specialbulletin
från

PEDAGOGISK-
PSYKOLOGISKA
INSTITUTIONEN

LÄRARHÖGSKOLAN
FACK, 200 45 MALMÖ 23

pedagogisk dokumentation

Bierschenk, I.:

DATORBASERAD INNEHÅLLSANALYS: KODNINGSMANUAL

Nr 52 Juni 1977

DATORBASERAD INNEHÅLLSANALYS: KODNINGSMANUAL

Inger Bierschenk

Bierschenk, I. Datorbaserad innehållsanalys: Kodningsmanual. Pedagogisk dokumentation (Malmö: Lärarhögskolan), Nr 52, 1977.

Denna manual är ett resultat av empiriska studier av ett omfattande intervjumaterial och flera utprövningar. Den har utvecklats på basis av en modell för formalisering av text, som bygger på dels psykologiska dels lingvistiska antaganden. Mot denna bakgrund har en psykolingvistisk modell utvecklats, på vilken en datorbaserad innehållsanalysteknik byggts upp. Beroende på hur modellens enskilda komponenter används kan olika innehåll utvinnas ur en text. Manualen är uppbyggd i sekvenser, vilka återger de olika stegen i modellen, från datainsamling via kodning av analysenheter till kontroller av överensstämmelse mellan kodare. Kodningsreglerna, som är huvuddelen, exemplifieras med verbalt material ur intervjuer. Exempelsamling och tekniska anvisningar för hålkortsstansning ges stort utrymme. Kodningsreglerna har utvecklats på grundval av talspråkstext men kan med fördel också tillämpas på skriftspråksmaterial.

Nyckelord: Datorbaserad innehållsanalys, psykolingvistik, datalingvistik.

INNEHÅLL

Förord

Denna manual är ett resultat av metodiska studier inom ett forsknings-projekt vid lärarhögskolan i Malmö. Projektet har finansierats av skol-överstyrelsens anslag för pedagogiskt forsknings- och utvecklingsarbete. Metodstudien är en del inom projektet, som har haft till syfte att utveckla ett system för en datorbaserad innehållsanalys.

Manualen är tänkt att kunna användas av olika grupper som arbetar med mycket varierande textmaterial. Den är utvecklad för analys av verbala in-tervjusvar, inspelade på ljudband och nedskrivna. Den presenteras i avsnitt som jag ansett vara av värde för andra projekt eller enskilda personer, som tänker sig att utföra liknande arbeten. Därför följer beskrivningen olika ut-vecklingsfaser, från val av undersökningsmetod via en stegvis utveckling av ett regelsystem för kodning av texter till beräkning av överensstämmelse mellan kodare av färdigkodad text för datorbearbetning. I anslutning till kod-ningsreglerna ges exempel från autentiskt textmaterial, eftersom reglerna har tillkommit genom empiriska studier. Min strävan har varit att i möjligas-te mån utveckla och exemplifiera reglerna enligt frekvens, vilket samtidigt medfört att vissa problematiska fall inte getts utrymme. Det kan tänkas att sådana fall uppstår beroende på materialets utseende och ämnesområdet i-fråga och att de bäst löses av den enskilde, antingen med hjälp av det regel-system som ges här eller genom egna utvecklingar.

För övrigt är manualen uppbyggd så att delar av den, som användaren tillfälligt vill arbeta med, kan särskiljas, t ex bilagorna 1-2. Bilagan med testmaterial bör inte finnas till hands för kodare.

Bakgrunden till utvecklingen av regelsystemet, som är den huvudsakliga delen i detta arbete, har redovisats i flera skrifter. En kortfattad presenta-tion av den teoretiska diskussion som ligger till grund för analyssystemet ges här i form av ett kapitel 0. Där redovisas också vilka analyser som hittills gjorts och vad en innehållsanalys på vårt sätt innebär (se också kap 1.2).

Till en presumtiv användare vill jag särskilt framhålla att den digra regelsamlingen inte nödvändigtvis måste användas i sin helhet. I kapitel 8 ges några exempel på vilka frågor man kan ställa sig till ett kodat material och vilka instruktioner en dator kan få för att ge svaren. Där kan det säkert gå att hitta egna idéer. Vilka koder som kan bli relevanta i ett visst samman-hang måste den enskilde användaren själv specificera med utgångspunkt i den egna analysens syfte.

Lingvistiska och psykologiska forskningsresultat av värde för mitt arbete har varit många. Att hänvisa till alla dessa på olika ställen i texten skulle göra läsningen och framförallt användningen av manualen onödigt svår. Jag har valt att ge en förteckning över den litteratur som haft betydelse för utarbetande av själva kodningsreglerna med tillhörande efterarbeten. Referenserna ges i sista kapitlet. Litteratur för de tidigare utvecklingsfaserna återfinns i de rapporter som anges i kapitel 0.

Till sist vill jag säga att detta arbete inte hade kunnat utföras utan att andra personer varit inblandade. Den som ytterst varit ansvarig för projektets uppläggning och genomförande är docent Bernhard Bierschenk. Hans omfattande psykologiska beläsenhet och kunnande har gett analyssystemet den nödvändiga bakgrunden och substansen. Jag vill tacka honom för att han uppmuntrat och stött mitt arbete och för att han delgett mig mycket av sina erfarenheter i olika sammanhang.

Fil kand Marjanna Berg har varit till stor hjälp i såväl kodningsarbete som överensstämmelseberäkningar. Andra personer har villigt hjälpt till vid olika kodningsförsök och på så sätt bidragit till att reglerna och inskolningsförslaget kunnat utarbetas.

Fil kand Leif Robertsson vill jag tacka för det intresse han visat för vårt material och för den intensiva arbetsinsats han gjort vid programutveckling och övrig datateknisk service. I detta sammanhang vill jag också nämna vår stansoperatris Ingegerd Johansson, som lagt ner stor möda och omsorg för att överföra ett omfattande material på hålkort.

Fröken Marianne Nyman och fru Anna-Lisa Gustavsson har väsentligen bidragit genom noggranna utskrifter av manuskriptet.

Flera personer har i mer eller mindre hög grad och i olika skeden läst manuskript och tidigare skrifter i anslutning till detta arbete. Jag är tacksam för de synpunkter som kommit. Vid utformningen av det slutliga manuskriptet har jag mottagit några råd av redaktionell art från professor Åke Bjerstedt och jag är tacksam för att han låtit manualen ingå i en av institutionens skriftserier. Innan manualen slutgiltigt färdigställdes har professor Sture Allén bidragit med värdefulla synpunkter av lingvistisk karaktär och jag vill gärna tacka honom för det intresse han visat. För såväl utformning som innehåll ansvarar jag naturligtvis helt och hållet själv.

Malmö i juni
I B

0. DATORBASERAD INNEHÅLLSANALYS: ANACONDA

Inom beteendevetenskaplig forskning har använts och används många olika innehållsanalystekniker, vilket innebär att beteendevetare kan anses vara väl förtrogna med såväl de teoretiska som de metodiska och tekniska eller praktiska problem som användningen av klassiska innehållsanalystekniker medför. En vetenskapligt genomförd innehållsanalys innebär att forskaren, oberoende av ett visst bestämt resultat, måste kunna redogöra för den valda metodens reliabilitet och validitet. En omsorgsfull, objektiv, reliabel och valid analys av text är emellertid mycket tidskrävande, och omfattande textanalyser kräver utveckling av mekaniserade eller automatiserade rutiner. I och med att datorer kan användas för ihågkommande och logisk selektion, hoppas vi kunna utveckla en datorbaserad innehållsanalysmetod som kännetecknas av större objektivitet och flexibilitet än vad som är fallet hos de klassiska innehållsanalysteknikerna.

Syftet med den här översikten är (1) att underlätta för läsaren förståelsen av vår metod för en analys av concept genom datorbearbetning (ANACONDA) och (2) att presentera ledtrådar, som vi hoppas ska ge en tillräckligt bra bild av de enskilda stegen i den kumulativa processen att utveckla en teori och en metod för att studera komplexa psykologiska fenomen, som t ex att meddela sig via symboler.

0.1 Problemställning

Att utveckla idéer, formulera och lösa problem är beteenden som är intimt förknippade med människans specifika förmåga att uttrycka sig verbalt. Försök att explorera och studera inte bara problemlösningsbeteenden utan även de komplexa psykologiska processerna bakom sådana beteenden sker inte bara inom psykologin utan sedan ett 20-tal år tillbaka inom flera andra vetenskapsgrenar, t ex matematik, artificiell intelligens, informationsbehandling och kvantitativ lingvistik. Det som förenar forskningsinsatserna inom de olika disciplinerna är målsättningen att utröna och göra synliga osynliga symboliska processer och mekanismer. Istället för att med hjälp av enkla matematiska formler försöka beskriva sådana kognitiva operationer, försöker forskare inom de nämnda områdena att utveckla datorprogram för att beskriva dessa och testa modeller och teorier om hur komplexa psykologiska strukturer ser ut och hur processer utvecklar sig.

0.2 Innehålls- och kontextorienterade teorier

Syftet med att försöka utveckla en teori om innebörden i meddelanden som män-

niskor kommunicerar sinsemellan är att öka vår förståelse av de kognitiva strukturer som antas ligga till grund för människans verbala uttryck. Genom att steg för steg utveckla och kontinuerligt utpröva våra antaganden försöker vi bestämma deras hållbarhet. Den fråga som vi har ställt oss är denna:

KAN VI MED HJÄLP AV NUMERISK ANALYS OCH GENOM KVANTITATIV BESTÄMNING IDENTIFIERA OCH KATEGORISERA KOGNITIVA STRUKTURER I VERBALA DATA, T EX INTERVJUTEXTER?

Mot bakgrund av de sista årens vetenskapliga debatt om processforskning och de påtagliga begränsningar som olika skattningsscheman som datainsamlingsmetod har, borde ett framgångsrikt genomförande av den uppgift vi angripit vara ett betydelsefullt bidrag till de forskningsmetoder som står till buds för samhällsvetare.

0.3 Guide till metodutvecklingen

Generellt om metod och modell. Skriven eller talad text kännetecknas av stor komplexitet och av att den typ av information som skall utvinnas ur ett material sällan eller aldrig är samlad på ett enda ställe i texten. Om strukturella relationer ändå skall kunna framträda krävs det att man preparerar texten på basis av vissa grundantaganden. I Bierschenk & Bierschenk (1976) presenteras ett flödesschema som anger de enskilda stegen i utvecklingen av den metod för datorbaserad innehållsanalys som vi föreslår. Dessutom presenteras den psykolingvistiska modell som ligger till grund för vår textanalys.

Nedan presenteras de olika faserna i kronologisk ordning.

Grundmaterial. Vi har avgränsat vårt empiriska material till att tills vidare gälla "Informationssökning, dokumentation och forskningsplanering för skolans F&U-arbete". Uppläggningen och genomförandet av en intervjustudie kring detta redovisas i B. Bierschenk (1974 a). I denna rapport utvärderas också skattningsskalor som intervjupersonerna konfronterades med under intervjun.

Impressionistisk analys. Ett sätt att utvärdera intervjutexter är att använda en impressionistisk innehållsanalys. Den bygger på intuition, inlevelse och intryck, som innebär att tolkningen bygger på subjektivt funna analysresultat. En sådan analys föreligger i Annerblom (1974).

Datorbaserad analys. En modell för sökning efter information i intervjutexter, en kortfattad beskrivning av preliminära kodningsregler och några empiriska resultat från utprövning av manuell tilldelning av koder på intervjutext presenteras i B. Bierschenk (1974 b).

Konstruktion av ett regelsystem. Förutsättningarna för att redan utvecklade metoder skall kunna tillämpas är ofta knutna till materialets utseende. De analysförsök som redovisats i litteraturen och som är av intresse för vår analys är utvecklade med skriven text som bas. Men eftersom vårt material är talspråkstext (utskriven från ljudband) som när den yttrades endast var avsedd för en person, intervjuaren, blev det nödvändigt att bygga upp ett eget system av regler och koder. En preliminär manual och några utprövningsresultat presenteras i I. Bierschenk (1974).

Tillförlitlighet i manuell kodning. Genom begreppet "datorbaserad innehållsana-

lys" vill vi markera att vi inte avser att utveckla en metod för automatisk textanalys. Detta betyder samtidigt att grundmaterialet först måste kodas innan maskinella bearbetningar av olika slag kan ta vid. Med vilken framgång två av varandra oberoende kodare har kunnat tillämpa olika kodningsregler på ett likartat sätt redovisas i Berg (1974).

Representation av manifesta språkstrukturer. Datorbaserad innehållsanalys börjar bli mer och mer använt och användbart internationellt. Kravet på program och tekniker som är avpassade för olika problemområden gör sig mer och mer gällande. Ett första steg i utvecklingen av en metod för analys av text kräver en modell för språklig representation. En startpunkt för vårt sätt att representera språkliga data utgjorde Schank's "Conceptual Dependency Theory", som i allt väsentligt bygger på kasusteori. På vilket sätt vi har utvecklat vår representationsmodell, på vilket sätt inmatning sker, hur identifikation specificeras i kodningen och hur vi har byggt upp vår lexikonbas redovisas i I. Bierschenk (1975).

Teoretiska och psykometriska problem. Att försöka kartlägga vad som i forskningslitteraturen egentligen means med en innehållsanalys är mycket svårt. Till grund för varje innehållsanalysteknik ligger nämligen ett specifikt sätt att betrakta innehållet i ett meddelande. En innehållsanalys förutsätter att vi kan definiera analysens mätobjekt, dvs det som skall mätas och räknas. Analysens enhet i vår psykolingvistiska modell går tillbaka på det välkända paradigmet Agent-aktion-Objekt (mål). Komponenterna definieras kortfattat nedan:

Analysens paradigm 1. Agent. Aktionscentrum eller målsökande entitet som använder sig av resurser för att realisera sina mål. Som agent fungerar personer, grupper, organisationer eller abstraktioner. 2. Aktion. En handling som utförs av en agent i syfte att realisera mål. Handlingen definierar innebörden i AaO-paradigmet. 3. Objekt. Allt som en handling kan vara riktad mot eller utföras med.

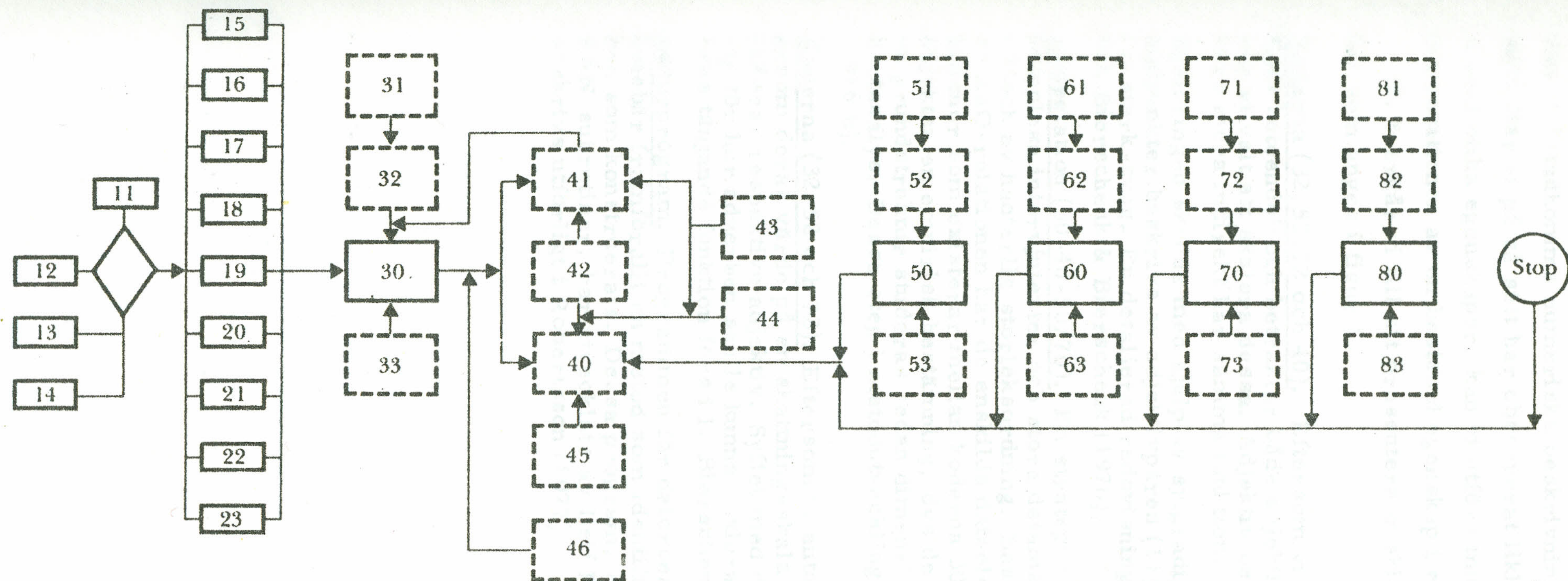
Genom AaO-paradigmet avgränsas de komponenter som utgör ett naturligt sammanhang, dvs en observerbar sats. Med detta avses den grundläggande formen för ett påstående som uttrycks genom substantiv$_1$ -verb-substantiv$_2$-relationen. Medan agent och objekt (substantiv) specificeras genom attribut (adjektivfraser), anges genom verbet relationen mellan substantiven, dvs handlingar, händelser eller tillstånd. Ordningen mellan dessa grundelement anges genom syntax. Genom ett lexikon och ett regelsystem hoppas vi kunna konstruera teorier och modeller som kan användas för att beskriva händelser.

Med hjälp av de lingvistiska element som finns i flödesschemat i figur 1 skall vi visa dels hur vi bygger upp concept i en given kontext, dels på vilket sätt vi avser att numeriskt beskriva texten och göra kvantitativa analyser. Hur vi utnyttjar dessa möjligheter beskrivs i B. Bierschenk (1976).

0.4 Numerisk beskrivning och kvantitativa analyser

Att bygga upp concept i en given kontext förutsätter ett regelsystem som anger hur olika element skall kopplas med varandra och i vilken ordning detta skall ske. Vi antar t ex att de relationer som finns mellan substantiv och adjektiv och mellan substantiv och verb återspeglar sådana relationer som kopplar empiriska fenomen med varandra. Om vi t ex vill ange att en uppsättning substantiv modifieras av en uppsättning adjektiv så kan vi formalisera denna relation.

Vidare antar vi att substantiv får sitt empiriska innehåll genom adjektiv och/eller verb som de är kopplade med. Genom att skalera adjektiv och verb



11 Informationskälla
12 Negation
13 Tid
14 Modalitet
15 Villkor
16 Orsak
17 Medgivande
18 Avsikt
19 Disjunktion
20 Komparation

21 Interrogation
22 Antagande
23 Önskan
30 Agent
31 Bestämning
32 Beskrivning
33 Klassificering
40 Händelse/tillstånd
41 Kopula
42 Satsmodifierare

43 Tidsbestämning
44 Platsbestämning
45 Handlings-/ tillståndsmodifierare
46 Omständighet
50 Objekt
51 Bestämning
52 Beskrivning
53 Klassificering
60 Riktning/lokalisering

61 Bestämning
62 Beskrivning
63 Klassificering
70 Mottagare
71 Bestämning
72 Beskrivning
73 Klassificering
80 Instrument
81 Bestämning
82 Beskrivning
83 Klassificering

Figur 1. Programflödesplan för en analys av text och uppbyggande av concept

kan vi åstadkomma numeriska beskrivningar och kvantitativa analyser av text. När vi på så sätt har observerat likheter eller kovariationer och definierat olika egenskaper, kan vi utföra multivariata analyser för att bestämma positionen av en bestämd egenskap i ett antal latenta dimensioner.

Nu övergår vi till att presentera enskilda analyser utifrån den kodstruktur som anges i figur 1.

Koderna (32,52,72 och 40). Eftersom vi i vår analys tar hänsyn till "syntaktiskt beteende" och betraktar både adjektiv och verb som beskrivande concept, har vi valt att skalera dessa. Adjektiv beskriver ett substantiv direkt, medan verb mera indirekt har samma funktion.

Skaleringen skedde med hjälp av sjugradiga skattningsskalor, vars bipolära ändpunkter beskrivs av adjektivparen (1) positiv-negativ, (2) aktiv-passiv och (3) stark-svag. En detaljerad redovisning av detta finns i B. Bierschenk (1976 a) och Bierschenk & Bierschenk (1976).

Kodrelation (30-40-50/70). En strategi för att övervinna de begränsningar som komplexa intervjutexter och stora datamängder medför är att bryta ned dessa i block av hanterlig storleksordning. Baserat på enkla numeriska beskrivningar av AaO-relationen har de enskilda dataelementen kondenserats. De logiska relationer som existerar mellan koderna 30-40-50(70) har utnyttjats för att ge objekten en empirisk bestämning, dvs de omvandlas till concept. På basis av en grundstruktur studeras sedan dimensionaliteten med hjälp av en diskriminantanalys. Dessa steg i metodutvecklingen beskrivs i detalj i B. Bierschenk (1976 b).

Koderna (32,52 och 72). Eftersom vi antar att adjektiv beskriver substantiven genom deras värde på en skattningsskala, har vi studerat de adverb som modifierar respektive adjektiv. Syftet med denna studie var att utveckla en princip för hur adverben skulle kunna skaleras. En principdiskussion kring adverbens tänjande funktion förs i I. Bierschenk (1977).

Datorprogram. Programmen för datorbearbetning enligt ANACONDA-tekniken innebär framförallt en metod som identifierar ord i texten med ord i de lexikon som konstruerats. Dessa program, som bygger på NU-Algol och FORTRAN-subrutiner, har utvecklats av Leif Robertsson vid Lunds Datacentral och beskrivs utförligt i Robertsson (1977).

0.5 Referenser till ANACONDA

Annerblom, M-L. En impressionistisk innehållsanalys av intervjuer med forskare på pedagogiska institutioner i Sverige. Pedagogisk-psykologiska problem, Nr 255, 1974.

Berg, M. Reliabilitetsprövning av en metod för innehållsanalys av intervjutext. Testkonstruktion och testdata, Nr 26, 1974.

Bierschenk, B. Perception, strukturering och precisering av pedagogiska och psykologiska forskningsproblem på pedagogiska institutioner i Sverige. Pedagogisk-psykologiska problem, Nr 254, 1974 (a).

Bierschenk, B. A computer-based content analysis of interview data: Some problems in the construction and application of coding rules. Didakometry, No. 45, 1974 (b).

Bierschenk, B. Teoretiska och psykometriska problem vid en datorbaserad innehållsanalys av intervjutext. Pedagogisk-psykologiska problem, Nr 287, 1976 (a)

Bierschenk, B. En datorbaserad innehållsanalys av intervjutext: Numerisk beskrivning och multivariat analys. Pedagogisk-psykologiska problem, Nr 307, 1976 (b).

Bierschenk, B. & Bierschenk, I. A system for a computer-based content analysis of interview data. (Studia Psychologica et Paedagogica, 32.) Lund: Gleerup, 1976.

Bierschenk, I. Konstruktion av ett regelsystem för en datorbaserad innehållsanalys av intervjutext: Preliminärmanual och några utprövningsresultat. Testkonstruktion och testdata, Nr 25, 1974.

Bierschenk, I. Datorbaserad innehållsanalys: Teoretiska och praktiska överväganden. Pedagogisk-psykologiska problem, Nr 283, 1975.

Bierschenk, I. En studie av modifierande ords kontextuella funktion: En principdiskussion och några kvantifieringsförsök. Pedagogisk-psykologiska problem, Nr 314, 1977.

Robertsson, L. Datorprogram till ANACONDA - en teknik för textanalys. Pedagogisk-psykologiska problem, Nr 317, 1977.

1. METOD, MATERIAL OCH MODELL

Det finns många vägar att gå för att samla in information. Vilket sätt man väljer är beroende av vilka frågor man vill ha svar på och från vilken källa man anser sig få de bästa svaren. Praktiska överväganden, som t ex tid, pengar eller geografiskt avstånd till källan, får ofta avgöra hur informationsinsamlingen kan ske.

Om man vill veta vad andra människor anser om ett problem som man ställer sig frågor kring, torde det bästa sättet att samla in den informationen vara att fråga de människor det gäller. Metoden man väljer kan vara mer eller mindre sofistikerad. Två vanliga tillvägagångssätt är enkäter, som skriftligen skall besvaras, och intervjuer, som är muntliga. Vid enkäter löper man risken att få otillräcklig information dels beroende på hur frågorna är ställda, dels beroende på hur svararen har kunnat arbeta sig igenom frågorna. Vid intervjuer är dessa risker mindre, eftersom intervjuaren kan be om förtydliganden och dessutom kan missförstånd vid utformningen av frågorna direkt rättas till. Men vilka tillvägagångssätt man än väljer av dessa två, är det svårt att undvika att få bristande information, om man håller sig enbart till fråga-svar-tekniken. Ju fler tillfällen man ger den tillfrågade att utveckla sina tankegångar, desto mer information kan man räkna med att få. Diskussioner mellan kontrahenterna med stöd av ett intervjuformulär är ett bättre alternativ. De i förväg bestämda frågorna ger struktur åt intervjun.

Om man väljer det senare alternativet och samtidigt är angelägen om att inte gå miste om vad som sägs, behöver man ta ställning till på vilket sätt man kan samla in materialet. Det smidigaste sättet är f n att använda bandspelare. Den som leder intervjun är då inte hela tiden tvungen att föra anteckningar utan kan ägna sig helt åt intervjupersonen.

Inom valet av metod bör också diskuteras valet av undersökningsobjekt, dvs hur många människor av en särskild kategori som ska intervjuas. Kan man arbeta med totalpopulation eller måste ett urvalsförfarande tillämpas och hur sker detta i så fall?

Sådana och liknande överväganden bestämmer hur undersökningsmaterialet (data) är beskaffat. Ett frågeformulär består av ifyllda frågor, där svaren samlats in i form av t ex kryssmarkeringar eller verbala svar. En bandupptagning är en talad text, mer eller mindre välformulerad och välstrukturerad. Det senare materialet behöver bearbetas på olika sätt innan det blir hanterbart. Är materialet stort bör banden först skrivas ut. Denna utskrift kan ske på flera sätt, som i sin tur beror på vad man vill undersöka, t ex dialekter, där intonation och fonologi spelar stor roll, eller talspråkets syntax och ordförråd, där

meningsbyggnad och morfologi spelar en stor roll, osv. Om man således är intresserad av hur någonting sägs, avgör detta på vilket sätt utskriften ska ske. Av betydelse är då att den som skriver (transkriberar) texten har nödvändig lingvistisk skolning. Är man däremot intresserad av vad som meddelas i texten och att detta "innehåll" så enkelt som möjligt ska kunna läsas, kanske man väljer att skriva texten i en form som liknar vanlig skriven text.

Den nedskrivna texten utgör råmaterialet på vilket en modell för vidare bearbetning appliceras. Modellen styr utformningen av regler och anvisningar för vad som är relevanta data i undersökningen och på vilket sätt dessa särskiljs från irrelevanta och kodas.

De ställningstaganden vi har gjort angående metod, material och modell är följande:

1. Intervjuformen har varit datainsamlingsmetod, där två tekniker använts, fråge - och intervjuformulär.
2. Populationen har varit 40 forskare på pedagogiska och pedagogisk-psykologiska institutioner i Sverige, som valts ut via slumpförfarande och enligt definitionen har "minst licentiatexamen".
3. Materialet har insamlats dels via frågeformulär med olika svarstyper, dels via bandupptagningar av all talad text.
4. Texten har skrivits ner på A4-ark. Principen har varit att den skall kunna läsas som skriftspråkstext, men skall i övrigt vara så autentisk som möjligt, dvs all hörbar text skrivs ut, även om sammanhangen verkar ofullständiga.
5. Modellen är en s k diskurs-modell, som innebär att information definieras med hjälp av lingvistiska relationer.

1.2 Realisering av modellen

När modellen är bestämd gäller det att bestämma vilken innehållsanalysteknik som kan vara lämplig för att realisera de intentioner man har med sitt modellval.

För det första har mängden av insamlade data betydelse. Vi har valt att använda oss av datorn, dels därför att det skulle ta alldeles för lång tid att göra en omfattande manuell analys av vårt stora material, dels därför att den information vi vill ha kan vara svår att strukturera i och med att svaren på frågorna inte alltid finns där man kan vänta sig dem.

För det andra styr modellen hur de olika komponenterna skall bearbetas. Vår s k diskursmodell innebär att de lingvistiska relationerna måste kunna framträda. De ska dessutom vara basen för den statistiska bearbetning som ska ske via datorn. Vi realiserar vår modell genom följande teknik för innehållsanalys:

1. Adjektiv och verb antas ge substantiviska element sitt empiriskt specificerade innehåll. Adjektiv och verb urskiljs därför.

2. Adjektiv och verb skaleras i Osgoods tre huvuddimensioner (positiv-negativ, aktiv-passiv, stark-svag) och deras vikter återförs till texten, dvs de kopplas till sina huvudord och ger dem "innehåll".

Den bakomliggande teorin är i korthet att såväl adjektiv som verb "uppgår" i sina huvudord, dvs de utgör egenskaper (här i tre dimensioner), som via en funktion härleds till sina respektive grupper.

Denna teknik möjliggör statistiska bearbetningar av olika slag, som strukturerar texten. Det är t ex möjligt att sammanfatta "innehåll" i stora textmängder genom en klusteranalys, där olika begrepp via t ex en matematisk funktion sammanslås och där grupper av begrepp fungerar på olika sätt i förhållande till varandra; de struktureras konnektivt, disjunktivt, kausalt, etc. Men tekniken förutsätter att

3. strukturer inom enheter och relationer mellan enheter inom en sats, samt relationer mellan satser upprätthålls och att
4. lexikon konstrueras, vilkas innehåll är så kongruenta med modellens komponenter som möjligt.

Detta kan realiseras genom konstruktion av ett regelsystem, efter vilket texten struktureras i enheter och relationer. Det är detta regelsystem som den här manualen presenterar.

2. URVAL OCH SEGMENTERING AV TEXT

Vårt textmaterial består av intervjudata som spelats in på band, dvs texten innehåller både frågor, svar och diskussion mellan intervjuare och intervjuperson (diskurs). Om man, som i vårt fall, väljer att skriva ut all text, blir det nödvändigt att bestämma hur mycket av texten som ska analyseras och motivera sitt ställningstagande utifrån teoretiska eller praktiska överväganden.

Vi ämnar inte studera interaktionen mellan parterna i intervjun utan är endast intresserade av vilka tankar intervjupersonerna uttrycker. En första anvisning för preparering av den utskrivna texten blev därför:

1. Gå igenom hela texten och stryk eller markera på annat sätt alla partier som innehåller intervjuarens frågor och övrigt tal.

   I vissa fall kan intervjuares och intervjupersons tal flätas in i varandras. Om man utan svårighet kan utläsa vad intervjupersonen menar, medtages yttrandet såsom intervjupersonens tal.

   Ex: Intervjuaren: Så att, dom här problemen med betyg har du inte...
   Intervjupersonen: funderat så mycket på, nej

   Intervjupersonen antas mena: Dom här problemen med betyg har jag inte funderat så mycket på, nej.

   Parenteser används för att skilja ut reella yttranden från implicerade. Detta förfarande behandlas under kapitel 5 om suppleringar.

I detta slag av intervjumaterial har man att göra med tämligen ostrukturerad text. Man kan inte vara säker på att ett uttömmande svar på en fråga verkligen ges på det ställe i intervjun där det var avsett. En strukturering av textmassan innebär därför att texten gås igenom och passager klassificeras i enlighet med vilken frågeställning de kan hänföras till. Svar på en fråga kan alltså finnas i anslutning till frågan, men också i anslutning till en diskussion som uppstått vid en annan fråga, som inte ens behöver vara likartat ställd. Vi har i vårt arbete erfarit att det ofta förhåller sig så. Till förberedelsestadiet hör därför också:

2. Ordna svaren till frågorna så att de tillförs respektive fråga. Frågorna bör ha ett nummer, som hjälper till att identifiera texten enligt den struktur som frågorna ger. Ju flera frågor man har desto lämpligare kan det vara att klassificera dem under ett antal huvudområden.

Det urval och den indelning av materialet man i början gör är väsentligt för fortsättningen, eftersom man skapar kodbeteckningar utifrån detta. Våra intervjupersoner är 40 till antalet, varför texterna fick ett nummer från 1 till 40. Frågorna har sammanförts till 7 huvudområden, men är 52, och varje frågas nummer kvarstår i kodningen. För att kunna behandla texten enligt huvudområden har vi en nyckel, som anger vilka frågenummer som

i så fall hör samman.

Vad som är relevant text är sådant som sägs som svar på intervjufrågorna eller som diskuterar ämnesområdet. Vid inspelningstillfället kommer ofta sådan text med som inte berör detta utan som kan vara en följd av att man bläddrar i papper eller intresserar sig för bandspelaren o dyl.

3. Text som berör själva intervjusituationen oberoende av frågorna stryks

Ex: Intervjupersonen: Vilken vill du jag ska besvara först
Är bandet på nu
Kan du ta loss dom här bladen
Jäklar.... nu funkar inte pennan heller

Den text som nu återstår är den som vi ska applicera våra kodningsregler på, dvs själva analystexten. Den texten ska segmenteras så att vi får basenheter att arbeta med för kodning av data.

2.1 Analystext

Beroende på om en text från början är ett skrivet dokument eller en bandad talsituation som nedskrivits är den olika svår att läsa. Den senare typen innehåller i mer eller mindre hög grad korrigeringar, felsägningar, tvekljud och liknande segment. En skriven text är (orto)grafiskt indelad i enlighet med skriftspråkets normer, t ex genom interpunktion. Hur den talande skulle ha gjort denna indelning i meningar och satser om han skrivit texten kan vi inte med säkerhet veta. Den personal som skrivit ner våra taltexter har tyckt sig urskilja gränser i talet och därför markerat dessa med kommatecken och punkt. Denna interpunktion kan möjligen vara till hjälp för segmenteringen av texten, men behöver inte vara det. Om vi hade varit intresserade av talspråklig meningsbyggnad, t ex såsom den uttrycks i pausering och tonhöjd, skulle vi ha gett helt andra anvisningar för transkribering, vilket i sin tur fordrat särskilt skolad personal. Vårt intresse är vad som sägs och inte hur.

Vad som sägs är emellertid inte alltid sådant som vi haft för avsikt att få veta. Allt som sägs är inte av betydelse för undersökningen ens efter de första urvalen. Vi har två ställningstaganden att göra, nämligen att bestämma (1) vilka textdelar som kan utgå och (2) hur indelningen av den återstående texten i analysenheter ska ske.

Inom den valda text som är relevant för undersökningen stryks

4. ansatser till eller bekräftelser på yttranden, dvs sådan text som antyder att intervjupersonen kommer att svara på en fråga eller anser sig ha avslutat svaret.

Ex: Intervjupersonen: Få se nu vad jag kan komma ihåg
Jag ska ta ett exempel
Men du, en sak till...
Ja, det var väl det huvudsakliga
Så ser jag på problemet

5. korrigeringar av talaren

Ex: Intervjupersonen: Det var projektledaren som först påpekade det... tror jag... jo det var det
Min professor... nja förresten det var t f som fixade det där

I det första fallet stryks 'tror jag', eftersom talaren korrigerar sig omedelbart och blir säker på vad han sagt. Därför stryks också 'jo det var det', som bara bekräftar att han kommit ihåg rätt. I det andra fallet utgår 'min professor' och 'nja förresten'.

6. obegripliga fragment av yttranden, dvs ett eller flera ord som inte kan sättas in i textsammanhanget och därur få sin mening. Hit räknas också tvekljud, som 'ööhhh', 'äum' o dyl eller harklingar. Tolkbara ord, som 'tja', 'njae' o likn medtages (se kap 4. 9. 3).

Det har visat sig vid våra utprövningar att texterna, trots lika innehåll, tar olika lång tid att preparera på det här sättet. Antalet strykningar blir oftast inte så stort, framförallt inte ifråga om korrigeringar eller obegripligheter. Strykningar behöver inte bli något stort problem, om man håller sig till en generell regel, nämligen:

7. Stryk inte lättvindigt. Ta hellre med ett osäkert parti. Det går att skapa en kod där osäkra fragment kan ingå.

När man vid olika slag av analyser av språkliga utsagor försöker definiera sina enheter bör man uppmärksamma begreppet 'mening', bl a därför att det kan tolkas som engelskans "meaning"(betydelse), dvs skillnaden gäller syntaktisk (eng. "sentence") och semantisk mening. När det gäller talspråk har man dessutom det problemet att grafisk mening passar illa (och egentligen inte finns). Vi har inte funnit det nödvändigt att på det här stadiet försöka oss på nya och svårigenkännliga definitioner på 'mening'. Vi håller oss här till den traditionella, välkända och inarbetade betydelsen, som utgår från satsbegreppet, där en fullständig sats innehåller en subjektsdel och en predikatsdel. En syntaktisk mening kan vara ofullständig, beroende på att vi ibland uttrycker oss ekonomiskt, där alla nödvändiga delar i den syntaktiska strukturen inte behöver finnas med. När vi läser eller hör den, implicerar vi dessa utelämnade delar och gör i tanken meningen fullständig. Vi får följande definition på mening:

En mening är en språklig enhet som består av

a. en fullständig huvudsats
b. flera satser, varav en är huvudsats och en eller flera bisatser
c. ord eller fras som kan byggas ut till fullständig huvudsats

Denna definition är tillräcklig för segmenteringen av texten i analys-enheter. Meningarna blir sedan föremål för bearbetning enligt vilken teori man har för vad som ska betraktas som "fullständighet", dvs vilka delar som måste ingå i subjekts- respektive predikatsdelen. För detta redogörs i kapitel 3 (se även kapitel 5). Innan vi går över till dem ska här exemplifieras hur man kan segmentera analystexten. Texten är hämtad ur en av våra intervjuer. Vi ger först den autentiska utskriften och därefter samma text i segmenterat skick. Motiveringen ges efter rutan.

Ruta 1. Exempel på intervjutext före och efter segmentering

I: Skulle du kunna säga något om hur man borde utforma informationssökningen för att skapa ideala förhållanden för forskningsprocessen?

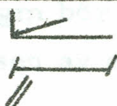
Ip: Jag är en väldigt dålig forskare när det gäller just sånt där. Jag är dj-ligt lite för att så att säga, jag tycker den här typen av problem är också, ja du motbevisar mig genom att peka på X:s uppsats, eller hennes avhandling, men mitt motargument är att du, ditt viktiga tips där, det var ett tips i tiden, hennes hade man säkerligen ändå fått koll på. Men jag är annars väldigt osystematisk i hela mitt sätt att söka och t ex jag har aldrig gått på någon sådan här referatsammanställningar någon gång, men däremot så, ibland har jag ägnat mig åt systematik när det gäller att leta igenom tidskrifter och sådant där, men den överger jag snabbt, medan jag däremot punktmässigt läser in mig på ett område eller en tidskrift väldigt väl, men det gör att jag, jag har dels utnyttjat väldigt lite den typen av sammanställningar, abstracts och sådant, och dels så har jag väldigt få rekommendationer. Vete fan om det inte är nån som har nån gång skrivit om det där att det är symtom på en intellektuell kris alltså, att man måste banta ner information för att kunna ta emot den alltså. Men jag kom inte riktigt ihåg vilken typ av ..., men det är alltså, för att man ska kunna använda information så måste man förstöra den alltså. ------

I: Skulle du ~~kunna säga~~ något om hur man borde utforma informations~~sökningen~~ för att skapa ideala förhållanden för forskningsprocessen?

Ip: //Jag är en väldigt dålig forskare när det gäller just sånt där.// ~~Jag är dj-ligt lite för att så att säga, jag tycker den här typen av problem är också, ja du motbevisar mig genom att peka på X:s uppsats, eller hennes avhandling, men mitt motargument är att du, ditt viktiga tips där, det var ett tips i tiden, hennes hade man säkerligen ändå fått koll på/~~ /Men jag är annars väldigt osystematisk i hela mitt sätt att söka//och t ex jag har aldrig gått på någon sådan här referatsammanställningar någon gång//men däremot så, ibland har jag ägnat mig åt systematik när det gäller att leta i-

Ruta 1.(forts)

genom tidskrifter och sådant där // men den överger jag snabbt, // medan jag däremot punktmässigt läser in mig på ett område eller en tidskrift väldigt väl, // ~~men det gör att jag~~ / jag har dels utnyttjat väldigt lite den typen av sammanställningar, abstracts och sådant, // och dels så har jag väldigt få rekommendationer. // Vete fan om det inte är nån som har nån gång skrivit om det där att det är symptom på en intellektuell kris alltså, att man måste banta ner information för att kunna ta emot den alltså. // ~~Men jag kom inte riktigt ihåg vilken typ av ..., men det är alltså~~, // för att man ska kunna använda information så måste man förstöra den alltså. // - - - - -



irrelevant analystext
strykning inom analystext
meningssegmentering //

Den första långa strykningen i analystexten består av en serie ofullständiga yttranden, som är svåra att göra fullständiga genom att man inte är säker på vad som saknas. I de första saknas möjligen predikatsdel. De följande berör något som intervjuaren har gett tips om och hör inte till frågan. De två sista strykningarna är också ofullständiga och hjälper inte upp framställningen om de finns kvar. Det återstår nio meningar, som ska analyseras vidare. Inom dem finns en hel del ord som skulle kunna utgå, främst talspråkliga småord, men de kvarstår för att vi inte ska behöva göra för stora ingrepp i personens sätt att uttrycka sig. Risken är annars att man stryker för mycket. Men intuitivt märker man vid genomläsning av dessa meningar att somliga ord anger ett huvudsakligt innehåll och att andra är mindre väsentliga för förståelsen. Men att bygga på intuition i detta sammanhang är inte tillräckligt. Kodningen måste systematiseras, så att kodaren med hjälp av regler kan urskilja väsentliga delar och så tillförlitligt som möjligt ange dessa enligt en vald modell. I nästa kapitel skall vi försöka beskriva hur identifikationen av textdelar går till och vad som menas med ett fullständigt kodat paradigm. När detta är klarlagt följer själva reglerna för tilldelningen av koderna och sådant som hör samman med identifieringen av enheter.

3. ANALYSENS PARADIGM: PRINCIPER

Vi har nu kommit så långt att vi avgränsat vår analystext och bestämt vad som är sats och mening. Den vanligaste principen för vad som är en fullständig sats kan uttryckas genom paradigmet AaO, som enklast betyder Agent-aktion-Objekt. Det betyder att någon handlar eller är föremål för en utsaga (A), att en handling eller händelse sägs ske (a) med eller mot ett objekt eller mål (O). En handling enligt paradigmet är alltid riktad (transitiv), explicit eller implicit.

Det är elementet a (verbet) som är viktigast att ha under kontroll när man ska identifiera satsens (paradigmets) ingående delar. Verbets betydelse anger vilket eller vilka objekt som måste finnas för att ett fullständigt paradigm ska urskiljas och kodas som analysenheter. Nedan ges några principer för hur ett fullständigt paradigm bör se ut och vilka olika typer av paradigm vi arbetar med. För förståelsen av det följande resonemanget hänvisas läsaren till figur 1.

3.1 Paradigmets horisontalled: Relationer mellan A⟷a⟷O

Som exempel på verbets betydelse vid identifiering av paradigmet kan vi ta verben 'skriva' och 'skicka'. Det första kräver en agent och ett objekt, det andra även ett mål (en mottagare):

Jag skriver rapporten	Jag skickade rapporten
A a O	A a O

Det första paradigmet är fullständigt. Det andra är ett fullt förståeligt meddelande. Om det har framgått vem som är mottagare, behöver det inte utsägas. Men vi som inte avser att göra automatisk analys kan inte förutsätta denna förståelse hos datorn, och måste alltså supplera (komplettera) den kunskapen:

Jag skickade rapporten (till SÖ)
A a O O (Mot)

"till SÖ" kan sägas vara det andra objektet som verbet 'skicka' kräver p g a den riktning som anges. O (Mot) betyder att objektet kallas mottagare.

Det finns andra verb som anger riktning, t ex 'resa'.

Jag reste till Stockholm
A a O (Mål)

Detta nya objekt, O (Mål), anger resandets slutstation. Det är inte mottagare av ett objekt, som 'rapporten' ovan. Skillnaden mellan O (Mot) och O (Mål) är denna: Agenten antas befinna sig på ett ställe och objektet flyttar sig i det första fallet. I det andra fallet är det agenten själv som förflyttar sig.

Finns det då inget "rent" objekt i det senare paradigmet? Jo, man kan säga att agenten har sig själv som objekt. Men det ligger i verbet 'resa', så det behöver vi inte identifiera särskilt. Däremot kan vi säga att det skulle ha kunnat finnas ett tredje objekt, som emellertid inte är nödvändigt, nämligen en uppgift om på vilket sätt resan skedde, t ex med bil, tåg eller flyg. När denna uppgift finns explicit (uttryckligen) kan man få denna utökning av paradigmet:

Jag reste till Stockholm [med flyg]
A a O (Mål) [O (Med)]

O (Med) anger det hjälpmedel man använt sig av. Man kan också kalla det teknik eller instrument, beroende på vilken utsaga eller text det gäller. Klammern omkring O (Med) anger att denna del i paradigmet är fakultativ, dvs förekommer frivilligt.

Vi har nu sett hur verbets karaktär styr vilket eller vilka slag av objekt som måste finnas. Det finns också ett annat kriterium för att identifiera riktningen, nämligen skillnaden mellan verbets aktiva och passiva form. Passiv form föreligger när verbet slutar på -s, brukar man säga (utom sådana verb, s k deponens, som i infinit form har -s, t ex 'att narras', 'att svettas'). Istället för s-form kan det också vara omskrivet med en form av verbet 'bliva'.

Rapporten skickade<u>s</u> till SÖ
Vi <u>blev</u> <u>utsatta</u> för kritik

Dessa båda satser är inte fullständiga enligt vårt analysparadigm. Agent saknas i båda fallen. I traditionell satsanalys har vi subjekt och predikat som kriterium på fullständighet, men här ser vi att det inte räcker. Vi måste komplettera:

Rapporten skickades till SÖ (av institutionen)
O a O (Mot) A

Vi blev utsatta för kritik (av lärarna)
O (Mot) a O A

Detta förfaringssätt anger riktningen i satsen. Verbet behöver i sig inte uttrycka att ett "riktningsobjekt" krävs (se 3.4 nedan).

3.2 <u>Paradigmets vertikalled: Relationer inom A, a, O</u>

Om nu verbet inte anger en rörelse mot ett mål, utan säger att någonting försiggår på ett ställe, så kan vi ta följande exempel:

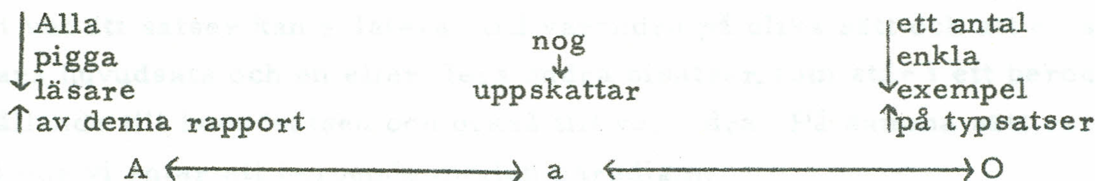
Jag bor på landet — Jag såg honom [på gatan]
A a O (Rum) — A a O [O (Rum)]

Samma struktur som det andra exemplet har detta, som anger tidpunkt:

Jag tänkte på dig [i eftermiddags]
A a O [O (Tid)]

O (Rum) anger plats för en handling, händelse eller tillstånd. O (Tid) anger tidpunkt för en handling, händelse eller tillstånd. Alla handlingar, händelser eller tillstånd förutsätter både rum och tid, och därför finns detta inte alltid angivet. Vid verbet 'bo' måste en sådan angivelse in för att paradigmet ska bli fullständigt. Rums- och tidsangivelser bör enligt vår uppfattning ligga utanför paradigmets horisontella led. Framförallt därför att verbets transitiva egenskaper inte framträder här. Däremot har vi tyckt det vara praktiskt att hänföra rums- och tidsbestämningar till verbet, såtillvida att de tilldelas samma huvudkod, eftersom man traditionellt brukar kalla sådana enheter för adverbial. Det finns också andra, som uttrycker sätt eller grad (sätts- resp gradadverbial) eller omständigheter. Ytterligare en typ uttrycker någonting generellt om hela satsen (satsadverbial). De här uppräknade typerna behandlas i kapitlet om kodningsregler.

Huvudkomponenterna A och olika O, som specificeras i sina relationer till varandra genom a, kan också bestämmas genom element som kopplas till dem och alltså räknas som beroende. De grupperar sig omkring huvudbegreppet enligt syntaktiska regler. Alla sådana element kan med ett gemensamt namn kallas attribut. De kan ha funktioner att kvantifiera, kvalificera, beskriva eller klassificera respektive A eller O. Några typiska beroenden visas genom denna uppställning horisontellt och vertikalt av en sats:



De ingående orden i blocken A och O har sina särskilda positioner i förhållande till varandra. Adjektiv står närmast huvudordet och beskriver eller kvalificerar det, före adjektivet bestäms huvudordet genom t ex en kvantifiering, efteråt kommer en klassificering, som ytterligare bestämmer eller avgränsar huvudordet. Även verbet har en beroende komponent, som här är ett satsadverb, som anger att hela satsen är ett (svagt) antagande.

Samtliga uppräknade typer av O som gjorts kan i princip ha samma slags bestämningar, som grupperar sig på samma sätt. Beroende på till vilken typ av O det hör får de identifierande koder i paradigmet. Utan att gå in på de olika koderna, kan den principen uttryckas så här:

Jag skickade den gula rapporten till biblioteket på LHM
A a O (beskr) O O (Mot) O (Mot_{klass})

Ytterligare ett förhållande bör klargöras, och det är sådana fall där vi inte har något Objekt utan där paradigmet fullständigas på så sätt att en kopula inträder istället för verb som en länk mellan Agent (begreppet sett här som "styrande" i påståendet) och någon annan del inom agentstrukturen:

Han är professor
Projektet kallas SÖK
SÖK-projektet är djupgående

Med liknande schematisk beskrivning som ovan kan ett par typexempel åskådliggöras så här:

Han ⟷ är ↑ professor

djupgående ↓ SÖK-projektet ⟷ är

Förbindelsen utgörs inte av objekt utan av en enhet som är beroende direkt till huvudordet och beskriver eller bestämmer huvudordet. Förhållandet måste uttryckas via 'är' för att en språklig sats ska bildas. Vi betraktar även en sådan sats som ett fullständigt paradigm, som till sig kan ha rums- och tidsbestämningar. Men vi särhåller sådana paradigm från de övriga genom att kopulan har annan kod än verben. Därigenom behöver man inte ta ställning till om huvudordet kan sägas vara agent eller inte. (A och O är till stor del en analysteknisk fråga.)

3.3 Relationer mellan flera paradigm: AaO⟷AaO

Vi vet att satser kan relateras till varandra på olika sätt och att en sats kan vara huvudsats och en eller flera andra bisatser, som står i ett beroendeförhållande till huvudsatsen och också till varandra. På samma sätt resonerar vi när vi antar ett beroende mellan paradigm:

Han reste till Köpenhamn för att delta i en konferens.

Det första paradigmet är fullständigt, det andra saknar agent, som återfinns i det första, och som vi alltså måste supplera in:

Han reste till Köpenhamn för att (han) delta i en konferens
A a O(Mål) ⟷ A a O

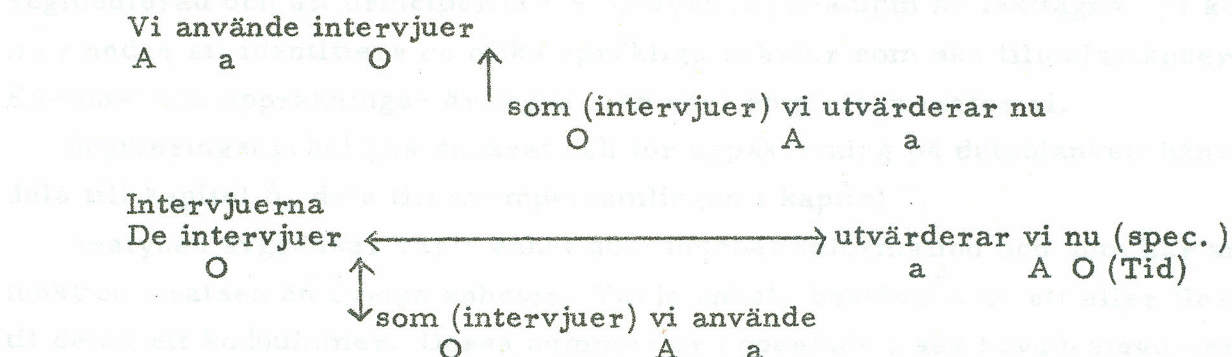
Uttrycket 'för att' relaterar den andra satsen till den första. Många sådana "kopplare" finns, som anger att satser binds ihop.

Ett annat sätt att relatera paradigm anges genom beroende inom paradigmet:

Vi använde intervjuer, som vi utvärderar nu.

De intervjuer som vi använde utvärderar vi nu.

Dessa exempel är två typiska relativkonstruktioner, som vi behandlar på ett likartat sätt:



Vilken konstruktion som än föreligger är objektet detsamma. Det första exemplet är en sammandragning av två satser, där den sista fått bilda bisats istället för att vara egen huvudsats (jfr "Vi använde intervjuer. Dessa utvärderar vi nu.") Det andra är exempel på hur en sats specificerar objektet i en överordnad sats. (Bestämningen är likvärdig med transformeringen "De av oss använda intervjuerna...".)

Med denna principiella översikt hoppas vi att grunden är lagd för förståelsen av de kodningsregler som utarbetats och som presenteras i nästa kapitel. En viktig del inom processen att identifiera och koda de olika textdelarna är att fullständiga dels paradigmet självt, dels beroenden och relationer inom och mellan paradigm. Detta förfaringssätt kallar vi supplering, vilket kommer att förtydligas i ett särskilt kapitel.

3.4 Kongruens med modell och teknik

Vi har sett att begreppet "riktning" (transitivitet) har den avgörande betydelsen vid identifieringen av de olika delarna. Att hålla isär de olika fallen av O är egentligen ett delprobelm, sett ur analystekniskt perspektiv, men de hjälper oss att i någon mån försöka renodla de typer av enheter som (ur klassificerings- eller kategoriell synpunkt) kan ingå i dessa O och som har betydelse för konstruktion av lexikonbaser. För denna strukturering är verbets funktion styrande.

Riktning ska ur analysens synpunkt också betraktas som ett funktionellt begrepp, nämligen som riktning i själva satsen. Överfört på analysens grundidé ses leden A respektive O som ingångar i en matris, där de två axlarna x och y inte får innehålla varierande funktionella element, men där individ - respektive variabelaxlarna för olika syften kan "byta plats".

Paradigmet AaO kan därför vändas till OaA, varvid de olika egenskaperna följer med. Verbet är nyckeln som anger den funktion som i varje enskilt fall ska tillämpas.

4. KODNING AV ENHETER, STRUKTURER OCH TEMAN

Tillämpningen av de regler som specificeras här förutsätter att analystexten är segmenterad och att principerna för analysens paradigm är iakttagna. Vi kommer nedan att identifiera de olika språkliga enheter som ska tilldelas koder. Exempel och uppräkningar är tagna från vårt empiriska material.

Suppleringsnyckel ges separat och för uppskrivning på datablankett hänvisas dels till kapitel 6, dels till exempelsamlingen i kapitel 7.

Analysen avgränsar varje enhet som innebär information och som har annan funktion i satsen än övriga enheter. Varje enhet, bestående av ett eller flera ord, tilldelas ett kodnummer. Dessa nummer är uppdelade i sex huvudkategorier, 3, 4, 5, 6, 7, 8. De motsvarar begreppen Agent, aktion, Objekt, Mål, Mottagare och Medel (se kap 3). Som underkategorier till dessa fungerar bestämningar av olika slag, som tilldelas siffror som omedelbart domineras av huvudkategorin: Agent får t ex siffran 30 och ett beskrivande adjektiv till agenten siffran 32, osv (se fig 1).

Presentationen av koderna och reglerna sker så att den högra spalten exemplifierar den vänstra spaltens kategoribeteckning.

4.1 Verb (Aktion) - kod 40

Definition: **En aktion är en riktad handling som utförs av (eller utgår från) en agent. Handlingen utförs mot explicita eller implicita objekt eller mål.**

1. finit verbform (presens eller preteritum)	sökte, avslutade, lyckas, har, fanns
2. finit hjälpverb + infinitiv eller supinum	skall överväga, hade diskuterat, skulle ha velat komma, kunde ha gjort
3. modala och andra hjälpverb + infinitiv: vilja, kunna, måste, skola, böra, må (månde), få, behöva, törs, (tör), torde, lär, verka, komma (att), bruka, börja etc.	Vissa projekt måste upphöra, bör starta, behöver ses över, torde misslyckas, verkade stå stilla, brukade läsa, började jobba
Vissa verb kan följas av objekt alternativt att-infinitiv. I dessa fall är enheten efter det finita verbet objekt.	Vi undvek att skriva (objekt) Han undvek personkontakter (objekt)
Vid t ex 'börja' inleds objektet med 'med', följt av 'att' som inleder hel sats. Hela satsen är då objekt. (se vidare kodning av 'att-satser', i exempelsamlingen, kap 7)	Jag började med att försöka etablera kontakter (objekt)

4. Verbbegreppet kan utvidgas om till verbet hör	
a) preposition eller adverb	söker upp, finns till, sätta ihop, dra till med, komma fram till, komma ihåg
b) reflexivum (fraser med 'sig')	lägga sig vinn om, ta sig samman, ta sig för, ta för sig, slå sig fram, sätta sig (ner), använder sig av, beter dig, informera mig
c) substantiv, som ger idiomatisk karaktär åt verbuttrycket	komma på kant med, lägga vikt vid, dra fördel av, göra lycka, ha effekt på, fatta beslut
Idiom: uttryck vars innebörd inte kan förutsägas på basis av de enskilda elementens betydelse(se vidare under kod 45)	hålla hus, visa tänder
5. a) verb + particip som är en omskrivning för passivum	blev använd, blir utnyttjad, skulle bli omtalad
b) övriga verb + particip	fått ... ökad, ha ... behandlade, finns inbyggd, gå förlorad
c) former av 'vara'+ particip eller adjektiv, följt av preposition (se vidare under predikativ)	vara intresserad av, var insatt i, var beroende av, var arg på
6. former av 'göra'+ substantiv där substantivets stam ensamt kan bilda verbet (se vidare under suppleringar)	göra sammanfattningar (=sammanfatta)

I princip kodas som verb sådana partiklar runt själva verbet som naturligast hör ihop med verbet och inte med ett efterföljande objekt. Man får i viss mån lita till språkkänslan när det gäller flerordsuttryck.

4.2 <u>Agent - kod 30</u>

<u>Definition:</u> Som Agent fungerar personer, grupper, organisationer eller abstraktioner. Agent är ett "aktionscentrum", dvs den enhet eller struktur som styr riktningen i satsen.

1. konkreta och abstrakta substantiv	projektet, universitetet, handledaren, SÖ, organisationskommittéer, bibliotek, system, Riksbankens Jubileumsfond, Torsten Husén, datalagen, seminarium
2. pronomen eller myckenhetsord som ersätter eller implicerar substantiv	han, det där, jag, vem, många, inga, andra,
3. självständiga adjektiv eller particip (som ej suppleras)	den tillförordnade, sakkunnige

4. infinitiv-fraser	Att bestämma sig var svårt. Det blev nödvändigt att informera om möjligheterna
5. underordnad sats	Att uppdraget gick till mig visar att man haft gott omdöme. Hur (på vilket sätt) det skulle gå till var diskutabelt. Om vi kommer igång är osäkert. Vem (vilka, vad) som styr är lätt påvisbart.

Lite hjälp för identifiering

Agenten är oftast liktydig med satsens subjekt. Utgå från var centrumet ligger, dvs varifrån satsens handlingsutsaga börjar. Om verbet står i passiv form står Agenten efter verbet, föregången främst av prepositionen 'av'. (Vid aktiv verbform kan Agenten stå både före och efter.)

4.2.1 Attribut framför huvudordet - kod 31

Definition: Ett attribut i denna ställning är en osjälvständig enhet inom agentkomplexet som har till uppgift att avgränsa och bestämma huvudordet, i första hand ifråga om mängd och ägande.

1. a) totalitet	alla, hela, fullständigt, totalt, inga
b) utpekande	denna, dessa, den här, mitt
c) possessiv/genitiv	dina, hans, projektets
d) kvantitet	många, flera, två, tusen, ytterligare
e) urval	övriga, första, de återstående, andra, annat
f) jämförelse	sådan, liknande, dylik
2. adjektiv som beskriver myckenhetsord	ett högt mått av intolerans, en hel (stor) hög med ärenden, en sabla massa strunt, ett litet antal böcker, nittio procent (av)
uttrycket kan omskrivas till myckenhetsord	en del, ett gäng (+ subst), mycket, massor av, många, några (få)
3. adverb som bestämmer myckenhetsord	ganska många, relativt få

4.2.2 Attribut framför huvudordet - kod 32

Definition: Ett attribut i denna ställning är en osjälvständig enhet inom agentkomplexet som har till uppgift att beskriva och värdera huvudordet. Detta attribut är först och främst adjektiv.

1. adjektiv	bra, dålig, bättre, sämre, stor, amerikansk, geografisk
2. particip	genomgripande, efterlängtad, reviderad, samlad, välskriven, välfungerande

3. adverb i vissa fall som fungerar adjektiviskt	en annorlunda bok, ett framåt projekt, gradvisa förbättringar
4. adverb + adjektiv, där adverbet uttrycker	
a) ringhet	mindre lyckad, något sämre
b) approximering	ganska bra, relativt bra, tämligen välskriven, rätt överskådlig, förhållandevis rikhaltig
c) exakthet	direkt farlig, precis vansinnig
d) myckenhet	mångfaldigt säkrare, mycket sämre, betydligt överdriven
e) emfas	verkligt relevant, riktigt bra, särskilt utvald, så vanlig
f) totalitet	fullständigt ny, allra äldst, alltför perfekt, definitivt kass, helt överspelad, rent för djävlig
g) extremhet	hemskt bra, väldigt dålig, enormt trevlig, kolossalt ambitiös
5. bestämningsord (enligt kod 31) som implicerar ett adjektiv, om detta klart framgår ur kontexten	sådan (relevant) information, liknande (irrelevanta) referenser

Anm. Till huvudordet inom denna kod hänförs också bestämd och obestämd artikel, samt likvärdigt indefinit pronomen (den, det, de, en, ett, någon, något, några).
Om 'några' kan tydas som antal ('några få') kodas det som 31!

4.2.3 Attribut efter huvudordet - kod 33

Definition: Ett attribut i denna ställning är en osjälvständig enhet inom agentkomplexet som har till uppgift att precisera huvudordet. Detta attribut uttrycker i första hand en klassificeringsaspekt.

Nedanstående redovisning sker så att de principiella typerna räknas upp i punkterna 1 - 6. De följande 3 är ojämförligt de vanligaste i vårt material och har här lagts in sist. Det har visat sig under våra utprövningar att den här koden har inneburit svårigheter. Därför exemplifierar vi med ett mycket större antal fall än för de andra bestämningarna.

1. adverb (uttrycker befintlighet, riktning, tid eller sätt)	Arbete hemma förekommer sällan. Undersökningar regionvis var att föredra. Tjänsterna dåförtiden var flera. Sneglingar mot USA är vanliga.
2. subjunktionsfras (uttrycker funktion)	Namnet som sådant är ointressant. Jag vill inte uttala mig i egenskap av delägare (äv. som delägare vill jag ...) (jfr punkt 6a nedan)
3. apposition	Han var personligen närvarande.

	Pettersson, institutionens vaktmästare. Kjell Härnqvist, professorn här.
4. framhävande tillägg (som ej är sats) (jfr kod 99, punkt 6)	som föregås av: bl a, till exempel, nämligen, alltså, det vill säga, etc
5. attributiv satsförkortning	Han gick, utom sig av oro ... (ovanligt i talspråkstexter)
6. när bestämningen är framförställd och inte kan vara kod 31, 32	
a) epitet	byrådirektör Svensson, professor Henrysson, skolrådet Marklund
behandlas likadant som relativkonstruktion	Svensson som är byrådirektör
b) genitivform som kan omskrivas till prepositionsuttryck	skolans utbyggnad (=utbyggnaden av skolan)
Genitivformen uttrycker inte ägande	Projektets tillkomst
7. infinitivfras	Idén att starta projektet. Förmågan att vara kritisk. Alternativ att välja mellan (på).
8. underordnad sats (inleds med 'att', frågande eller relativt pronomen/adverb)	Det faktum att vi låter oss styras ... Frågan vilka problem som skulle ventileras ... Vi ville inventera problem som lärarkandidater ställs inför.
9. fraser (ibland satser) inledda med preposition De vanligaste inledarna är: av, från, i, för, med, till, på, vid (om satsförkortningar inledda med 'med', se kap 5)	Resultatet av våra ansträngningar blev oväntat positiva. Behovet av bättre resurser är stort. Miljöer av typen Rosengård... Erfarenheter från arbetet med LGR-69... Referenser från ERIC datorbaserad... Utdrag från originalet... Anvisningarna i den bok som används... Boken i original... En kurs i bearbetning av data... Litteratur för självstudier... Rapporter för den större läsekretsen... Tid för att kunna skriva... Avsikten med rundresan... Kontakter med personer utanför projektet... Anledningen till skrivelsen var... Resan till konferensen... Beviset på att detta borde gjorts... Samlingarna på biblioteket... Forskningen vid lärarhögskolorna...

Inledarna kan också vara uttryck som: i anslutning till, i relation till, när det gäller

Problem när det gäller projektforskningen... Litteratur i anslutning till mitt område...

4. 3 Predikativ

Definition: Predikativ är en utsaga som sägs om agenten och som förbinder denna utsaga med huvudordet genom en kopula.

4. 3. 1 Kopula - kod 41

verb vid predikativ

a) vara, bliva, heta, kallas (traditionell predikatsfyllnad)

Han är docent. Dom kommer att bli (har blivit, etc) etablerade. Tidningen kallas (heter) Pippi.

b) liknande verb, som fyller funktionen av att vara predikativa: anses, tyckas, förefalla, verka, se ut, representera, m fl.

Hindren föreföll oöverstigliga. Dom anses omöjliga. Jag fungerade som projektledare. Jag satt som sakkunnig i flera år.

Det efterföljande ledet kan då vara

1. adjektiv, particip eller båda dera (kod 32)

Skrivelsen är bra formulerad. Den är lättläst, korrekt.

2. substantiv eller substantivfras enligt regler för agent (kod 30)

Uppgiften var att etablera kontakter. Idén var att starta en referensgrupp.

För att skilja på kod 33 resp. 30 kan man utgå från den "syntaktiska balansen" i meningen. Efter en kopula kan vi inte ha kod 33. Den måste stå direkt intill huvudordet.

Uppgiften att etablera kontakter var
30 33 41

Adjektivet är predikativ

svår
32

Adjektivet står som bestämning framför huvudordet. Att-frasen är predikativ.

Den svåra uppgiften var att etablera kontakter
32 30 41 30
30

4. 3. 2 Verb eller predikativ?

Vi utgår ifrån att skillnaden mellan adjektiv och verb ofta är obetydlig, de är båda "predikat". När vi ska särskilja dem har det med deras funktion att göra. Fraser med former av 'vara' är verb (se under 4. 1, punkt 5) när de följs av objekt. En indikation på att objekt följer är en preposition. Utan denna indikation kodar vi fraser som predikativ och själva kopulan får då kod 41.

a) verb vid något slag av objekt

Han var intresserad av rapporten
30 40 50

b) predikativ — Han var ledsen
30 41 32

(Obs dock kap 9.4.1.)

4.4 Objekt - kod 50

Definition: Objektet är den enhet som påverkas av handlingen i verbet.

Strukturen inom koden Objekt (50) ser likadan ut som agentstrukturen. Uppdelningen av de olika elementen i 51 - 52 - 50 - 53 sker därför enligt samma principer. I övrigt hänvisas till kapitel 3, samt kodningsexemplen.

Lite hjälp för identifiering:

När verbet står i passiv form finns objektet att söka först i satsen. Objektet är oftast liktydigt med "direkt objekt". Objektet förekommer också vid de utvidgade verben (se 4.1, punkt 4), där olika partiklar definierats som tillhörande verbet och inte Objektet.

4.5 Objekt (Mål) - kod 60

Definition: Objekt (Mål) anger en punkt från eller mot vilken en agent förflyttar sig.

Strukturen inom koden Objekt (Mål) (60) ser likadan ut som agentstrukturen. Uppdelningen av de olika elementen 61 - 62 - 60 - 63 sker därför enligt samma principer. I övrigt hänvisas till kapitel 3, samt kodningsexemplen.

Lite hjälp för identifiering:

Objekt (Mål) är en enhet som tar Objektets plats vid verb som uttrycker rörelse mot eller från något. Prepositionen hör till detta objekt och inte till verbuttrycket. De vanligaste prepositionerna eller adverben som implicerar detta objekt har visat sig vara: från, dit (suppleras), till.

4.6 Objekt (Mot) - kod 70

Definition: Objekt (Mot) anger mottagaren eller målet för en agents handling med ett objekt.

Strukturen inom koden Objekt (Mot) (70) ser likadan ut som agentstrukturen. Uppdelningen av de olika elementen 71 - 72 - 70 - 73 sker därför enligt samma principer. I övrigt hänvisas till kapitel 3, samt kodningsexemplen.

Lite hjälp för identifiering

Objekt (Mot) finns alltid som ett andra nödvändigt objekt, förutom Objektet (50) och kan därför liknas vid s k "indirekt objekt". Det anger en konkret eller ab-

strakt punkt dit ett objekt förflyttar sig. Prepositionen hör till detta objekt och inte till verbuttrycket. De vanligaste prepositionerna vid Objekt (Mot) är: till, för, åt. De vanligaste verben som tar objekten 50 och 70 är få, ge, ta, lova, säga, tala, diskutera, fråga, svara, be. (En lista ur vårt intervjumaterial ges i bil 1:3.)

4.7 Objekt (Med) - kod 80

Definition: Objekt (Med) anger medel, tekniker, etc. med vars hjälp en agent utför en handling.

Strukturen inom koden Objekt (Med) (80) ser likadan ut som agentstrukturen. Uppdelningen av de olika elementen 81 - 82 - 80 - 83 sker därför enligt samma principer. I övrigt hänvisas till kapitel 3, samt kodningsexemplen.

Lite hjälp för identifiering:
Objekt (Med) är inte lika vanligt förekommande som övriga objekt. Det är inte nödvändigt i paradigmet, vilket ofta medför att det står efter de övriga huvudbegreppen. Prepositioner som föregår Objekt (Med) är i första hand genom, med (hjälp av), via. Förutsättningen för att uttrycket kodas 80 vid 'med' är att substantivet inte kan göras om till adverb/adjektiv (t ex med datorundersökning, med lite hopmyglade medel, med hjälp av bibliografier) (jfr kod 45).

4.8 Adverbiella bestämningar

4.8.1 Satsmodifierare - kod 42

Definition: Satsmodifierare uttrycker hur säker eller osäker den talandes utsaga skall uppfattas.

1. a) ett-ordsadverbial som är allmänna, modala eller konjunktionella	ju, väl, alltså, nog, kanske, liksom, naturligtvis, egentligen, säkert, troligen
b) särskilda satsadverbial (se bil 1:1)	bara, också (=även), inte ens, redan, (inte) än, ännu, just (just idag), däremot, (igår däremot)
2. satsadverbialssats, som har samma funktion som ett-ordsadverbial	
a) uttrycker säkerhet, antagande, tvivel	Han (NN) har medverkat, skulle jag tro (=troligen, antagligen) Det är otvivelaktigt att jag kommer i tid (tvivelsutan, säkerligen) Hon (NN) har många järn i elden. Det är jag säker på.* Jag tvivlar på att det (X) går igenom.
b) uttrycker talarens åsikt eller ståndpunkt	Jag tycker att, tycker jag, för min del anser jag att, m fl.

3. negationer har egen kod — (se punkt 4.9.2)

Om satsen inte är negerad kodas 42 på negationsordet (se ex under 4.10, punkt 2)

* Obs! att satsadverbialssatsen även kan bilda egen huvudsats. När satsmodifierare uttrycker antagande eller osäkerhet kodas det även i satstemakolumn (se 4.10, punkt 12).

Lite hjälp för identifiering:

Här hänvisas till bilagorna 1:1 - 1:2, som ger en fyllig bild av den typ av uttryck som kan förekomma. Det utmärkande draget är att varken orden, uttrycken eller satserna säger något om det innehåll talaren vill förmedla.

4.8.2 Tidsbestämning - kod 43

Definition: Tidsbestämning anger inom satsens (verbets) betydelseområde tidsaspekten för handling, händelse eller tillstånd.

1. adverb eller -fraser som uttrycker	
a) avgränsbar eller specificerad tidpunkt	i morgon, igår eftermiddag, kl 12.00, första januari 1974, förra året, i slutet på månaden.
b) tidsadverb, som kan suppleras med specificerad tid	nu, då
c) tidsåtgång eller frekvens, även duration (fortgående)	i många timmar, dagligen, ofta, alltid, en gång i månaden, hela tiden, regelbundet, så länge, hittills
2. fraser som uttrycker förlopp, tillfällen, skeden (faser), duration (varaktighet)	under problemformuleringen, i ett seminarium, efter en treårsperiod, under tiden
3. hela satser	inledda med tidsadverb = då, när, medan, sedan, etc.

Lite hjälp för identifiering:

Förutom ett-ordsadverben inleds tidsfraserna oftast med prepositioner eller adverb, varav de vanligaste har visat sig vara: i, under, vid, på, för ... sedan, efter.

Tidsbestämningar är inte nödvändiga delar av paradigmet och är inte beroende av verbets eventuella riktning.

4.8.3 Rumsbestämning - kod 44

Definition: Rumsbestämning anger inom satsens (verbets) betydelseområde rums- eller platsaspekten för handling, händelse eller tillstånd.

1. substantivfraser som anger	
a) ett konkret avgränsat (geografiskt) område	på biblioteket, i det här landet (Sverige), inom Omsorgsstyrelsen, nere hos er i Malmö, uppe i Norrland, ute på Hermodsdal, vid psykologiska institutionen i Köpenhamn
b) en definierbar eller klassificerbar företeelse	i ett projekt, i vuxenutbildningsutredningen, i utrikisk forskning, i vår (svenska) vetenskapstradition, i tidskrifter
c) odefinierad bestämning av plats eller företeelse	på alla nivåer, på flera områden, på en viss punkt, i dagens samhälle, inom en viss ram
2. rumsadverb, som kan suppleras med specificerad plats.	här, där, (däruppe, härnere), etc.
3. hela satser	inledda med rumsadverb, oftast 'där'

Lite hjälp för identifiering:
Förutom ett-ordsadverben inleds dessa bestämningar med prepositioner, som här anges med sjunkande frekvens: i, på, inom, vid, hos, här på, ute i, uppe på, nere hos.
Rumsbestämningar är inte nödvändiga delar av paradigmet och är inte beroende av verbets eventuella riktning.

4.8.4 Handlingsmodifierare - kod 45

Definition: Handlingsmodifierare specificerar eller värderar en handlings utförande till sätt eller grad.

1. adverbiella adjektiv, ensamma eller modifierade	Det (X) sker automatiskt, det går fortare. Jag brukar redovisa det hela väldigt exakt. Jag söker intensivt, för det mesta manuellt. Jag önskar att litteratursökningen skulle göras mera systematiskt. Jag anstränger mig inte mycket, men kontinuerligt.
2. ett-ordsadverb	Vi kontaktar varann brevledes. Avrapporteringen har gjorts etappvis. Vi gör det gärna, personligen.
3. fraser, varav många har idiomatisk karaktär	för hand, hand i hand, helt och hållet, i bokform, i förväg, i detalj, på allvar, på direkten, i form av siffror, av en (ren) slump, enligt mitt system, på det där strikt logiska sättet, på lämpligt vis, till vissa delar
4. hela satser	på det viset att ...

Lite hjälp för identifiering:

Handlingsmodifierare av det här slaget är förknippade med verbet och är inte någon nödvändig del inom paradigmet. Dessa enheter kan skiljas från kod 80 främst därigenom att prepositionerna som inleder fraserna är andra: i, på (där frasen oftast avslutas med 'sätt', 'vis' och 'väg'), till. När prepositionen är med, omedelbart följd av substantiv som kan göras om till adverb/adjektiv kodas 45 (t ex med ointresse = ointresserat, med glädje = glatt). I annat fall kodas 80 (se denna kod).

Vid idiomatiska uttryck (enligt 4.1, punkt 4c), som inbegriper ett adjektiv, t ex 'lägga stor vikt vid', betraktas modifieraren som adverb, dvs 'lägga vikt vid' är verbet och 'stor' får kod 45.

4.8.5 Omständigheter - kod 46

Definition: Omständighetsuttryck anger ramarna för handlings, händelses eller tillstånds giltighet och är underordnade ett huvudords betydelse.

1. fraser, som anger	
a) 'i vilket avseende' och som inleds med idiomatiska uttryck	när det gäller, då det gäller, i samband med, med tanke på, i fråga om, med avseende på
suppleras i vissa fall	i detta fallet (design), i det här fallet (problempreciseringskrav), i den meningen (att göra bekantskaper)
b) förutsättningar och förbehåll	med den bakgrund vi har, med det mångsyssleri som finns, med hänsyn till de problem man bearbetar, i förhållande till problemställningen.
2. fraser eller ett-ordsuttryck som anger	
a) komparation	än, lika ... som, liksom, såsom, mer ... än, ännu ... än, ... som
b) orsak	beroende på, på grund av, för ... skull, tack vare, i och med, genom min doktorsavhandling
c) medgivande	trots ...
suppleras i vissa fall	ändå (trots prishöjning)
3. satser, inledda med subjunktioner, där endast det inledande ordet (frasen) får denna kod	
a) komparation	än att, än, liksom, såsom, så
b) orsak	eftersom, därför att, genom att, beroende på att, i och med att, på grund av att, för att, tack vare att, av den anledningen att, (då)

c) villkor	om, ifall, på villkor att, förutsatt att, i den mån, såvida, såvitt, under förutsättning att, bara (blott), m fl
d) följd/avsikt	för att, så att, så, för
e) medgivande	trots att, även om, fast, fastän, låt vara att, medan, samtidigt som, m fl
f) disjunktion	antingen ... eller, medan ... däremot visserligen ... men, å ena sidan ... å andra sidan
g) fråga	om (huruvida), ifall

Dessa satsinledare anges således samtliga med kod 46. När det gäller att särskilja dem från varandra till betydelsen används ett system av koder i särskilda temakolumner, vilket presenteras i kapitel 4.10.

Lite hjälp för identifiering:

Punkterna 2 och 3 torde inte innebära några svårigheter. Punkt 1 innehåller däremot exempel som antyder det något "luddiga" i den här koden. Ett råd är att se efter idiomatiska uttryck som inte passar in i någon annan kod. Uppstår ytterligare osäkerhet om enstaka småord, se kod 99.

4.9 Övriga enheter

4.9.1 Bindning mellan begrepp - kod ++

1. när ett 'och' eller liktydigt uttryck binder ihop två huvudord	Jag skrev rapporter och PM. 50 ++
2. när bindningen samtidigt förekommer i negerad sats.	Jag hade varken tid eller pengar. 50 ++ 50
	Vi hade inte vare sig lust eller ork. 50 ++ 50

4.9.2 Negationsord - kod --

1. minustecken används för olika grader av negationsord, som uttrycker

a) helt negerande	icke, inte, ej, aldrig
även substantiviska	ingen, ingenting
b) första leden i uttrycken (jfr bindingen ovan)	varken ... eller, vare sig ... eller, ... vare sig lust eller ork. -- ++
c) osäkert negerande	knappast, knappt, näppeligen
d) subjunktion	utan att,

Anm. Vi ser här 'eller' i första hand som en konjunktion, men innebörden (temat) i hela satsen behöver inte gå förlorat om man kodar även 'eller' med --. (Se vidare om temakodning, kap 4.10, punkt 2.)

4.9.3 "Innehållslösa" utfyllnads- eller formord - kod 99

Många av de s k ersättningsorden eller andra småord används ofta i talat språk utan att ha egentlig funktion. Man kan naturligtvis diskutera deras eventuella innehåll, men vi menar att de ur innehållslig synpunkt är ointressanta för vår analys.

1. adverb utan sin vanliga funktion ("talspråksord")	Februarinumret i år handlar om Principals and Vice Principals. Då har dom samlat allt som är skrivet. (då uttrycker ingen tidsrelation.)
	Nu gick det inte bra för den där utredningen. (nu anger ingen tid.)
	Sen vill jag gärna understryka ... (sen uttrycker endast att talaren fortsätter.)
	Efteråt, så blev vi bjudna till konferenser. (så kan vara ett talspråkligt uttryckssätt eller insätts för att göra texten mera lättflytande.)

Vi vill inte stryka dessa småord ur materialet, dels därför att många av dem skulle kunna ha betydelse vid närmare granskning (det finns risk för missbedömning), dels därför att vi så lite som möjligt vill göra intrång i intervjupersonernas sätt att uttrycka sig. Finns dessa småord med i utskriften tvingas kodaren tänka igenom betydelsen av dem och tilldela dem ett kodnummer. Kod 99 kan sägas vara en kontroll på att vissa "osäkra" ord har kommit med. Det är lätt att efter kodningen kontrollera om det i den "innehållslösa" koden finns sådana ord som kan kodas om, eftersom de kan identifieras.

2. formellt subjekt eller emfas	a. Det fanns många projekt på gång. b. Det var flera projekt som hade lämnat in ansökan om tilläggsanslag.
innebörden bibehålls, dvs	a. Många projekt fanns på gång... b. Flera projekt hade lämnat in...
3. s k opersonlig konstruktion	Det skrivs promemorior. Det jobbas här. Det regnar.
4. konjunktioner som inleder huvudsats ("länkord")	Vi har jobbat väldigt mycket som ett team och vi har inte gjort en klar avgränsning från början.
dvs	1. Vi har jobbat... 2. Och vi har inte
andra konjunktioner	men, för, utan, ty, så, fast

5. ja, jo, nej-svar	Nej jag tänker inte åka.	
Om svaret fortsätter med ja, jo, nej och liknande, förs svarsorden till denna mening som första ord.	Du är handledare? Jo. Jo (jag är handledare)	
även utan svarsinnebörd	Har du ett eller flera projekt? Ja man kan ju säga att det är flera, eftersom det består av tre avgränsade delar.	
Ord med samma tolkning	tja, 'njae', jaha då, o. likn.	
6. förkortningar och andra "småord" som inte går in under satsmodifierare (kod 42)	t ex, bl a, etc, alltså	
Om de förtydligar en bestämning tillhör de bestämningskoden (se kod 33, punkt 4)		

Lite hjälp för identifiering:

Den här koden tenderar att bli "slaskkod". Det är därför svårt att ange exakt vad den kan innehålla. I första hand bör man tänka på det och det-var-som enligt regel 2-3. Vanligtvis inga subjunktioner. Lite hjälp ger dessa ord, som i vårt material är de mest frekventa och som är lätta att känna igen mot bakgrund av reglerna:

det, så, som (enligt p. 3), och, men, är (enligt p. 3), då, ja, utan, nej, där, sen, för, var (enligt p. 3), nu, så att, dels.

4.10 Övergripande teman

De följande kodtyperna skiljer sig från de övriga. De anges inte per enhet utan per sats, vilket gör att koderna kan vara desamma (bestå av samma ensiffriga tal), men särskiljs genom att de specificeras på en viss given plats i datablanketten, se kap 6). Det vi här ska göra är att identifiera och exemplifiera dessa teman och ge dem en kodsiffra. För uppskrivning på datablankett hänvisas till kapitlen 6 och 7 samt bilaga 2:1. Presentationen sker här inte efter frekvens, utan efter kolumnernas ordning. (Kol 11-23 anges i mittspalten.)

	Kol		Kod
1. Källan till utsagan, om man vill hålla isär åsikter av intervjupersonen själv och refererade åsikter (säga, tycka, tro, etc)	11	Talaren själv anser	1
		Talaren refererar till någon annan	2
Vid "indirekt anföring":		Han säger att ...	(2)
		Han påstår att ...	(2)
Åsikten innefattas även av den talande		Han vet att ...	(1)

	Kol		Kod
Detta "tema" tillhör identifikationen av textavsnitt (se kap 6)			
2. Negationer uttrycks genom negationsord (se punkt 4.9.2)	12	Vanliga negationsord + varken ..., vare sig	1
		knappast m fl	2
		subjunktionssats - utan att...	3

Negerade satser kan vålla stora problem, eftersom man inte kan undvika de filosofiska begreppen 'falsk' och 'sann'. Vi försöker här att vara så formella som möjligt, för att undvika tolkningsolikheter. Några exempel:

a) "säker" negation (giltig)		
påstående	Jag är inte forskare	(1)
disjunktion	...varken bra eller dålig	(1)
komparation	X var inte bättre	(1)
medgivande	Trots att vi fick hjälp lyckades vi inte...	(1)
avsikt	Vi genomdrev planerna för att vi inte skulle framstå som ynkliga	(1)
följd	Han kom så sent att han inte kunde vara med.	(1)
orsak	Jag kom inte, eftersom jag inte hade adressen.	(1) (1)
b) villkor (giltigt)	Om jag inte lyssnar så får jag inget höra	(1) (1)
	Om vi hade kunnat undvika reaktionen så hade vi inte sagt något	(1)
(hypotetiskt)	Om du inte antar erbjudandet så får du inga fler chanser	(1) (1)
Und. a. frågekonstruktion (ej neg., se kod 42)	Jag undrade om det inte var värt en satsning	
b. antagande	Jag tror inte att det lönar sig	(1)
c. önskan	Jag skulle inte vilja att du refererade det här	(1)

	Kol		Kod
3. Tempus uttrycker verbets betydelse i tiden, sett från den talandes utgångspunkt	13	Utifrån talaren nuvarande	1
		Utifrån talaren förfluten	2
		Utifrån talaren kommande	3

De tre reglerna innebär att vi inte enbart utgår från den tempusbetecknande verbformen.	...och när jag kommer dit (2)
	så får jag uppleva detta... (2)
	Han blir nog docent (3)
Vissa former uttrycker i kontexten en annan tidsaspekt än vanligt	Vi hade bestämt oss för att vi skulle ta upp frågan (2)

	Kol		Kod
4. Modus modifierar huvudverbets betydelse	14		
1. indikativ		Uttrycker att en handling eller skeende framställs som verkligt, utan att ange talarens vilja, önskan, förmåga eller liknande.	1
2. imperativ		Uttrycker en uppmaning (som i skrift ofta anges med utropstecken).	2
3. konjunktiv		Uttrycker något icke-verkligt eller hypotetiskt.	
1) konjunktiv form		Verbformerna vore, finge, (andra är mycket sällsynta).	3
2) konjunktiv betydelse			
a) vid ´skulle´		skulle (+infinitiv) skulle ha (+supinum)	4
b) vid uttryck för önskan, villkor eller jämförelse		Jag skulle önska (vilja) att det fanns referenslistor Om det inte hade varit såna mängder, så hade man kanske fått en bättre överblick. Det verkar som om det hade varit till största delen amerikansk påverkan.	
c) modala hjälpverb som styr infinitiv		behöva, böra, få, kunna, måste, skall, torde, vilja, komma (att), verka,	
(även mindre vanliga)		låta, lär, må (månde, måtte), törs	
4. aspekt som uttrycks med vanliga hjälpverb, förutom temporala och modala		bruka, börja, sluta, tänka, försöka, hinna, orka, råka, våga	5

Lite hjälp för att identifiera konjunktiv

Ett bra kännetecken på konjunktiv (när den inte är explicit enligt punkt 1-2) är att satsens verb ofta har en tidsform som anger förfluten tid men uttrycker någon slags önskan i framtiden. Det är alltså fråga om något som i realiteten inte har ägt rum. Kontrollera i första hand önskan, villkor och jämförelser (av typen ´som om´).

Anm. Istället för att dela på de olika typerna av modus mer än vi har gjort, har vi i vårt kodsystem möjlighet att särhålla dem genom kombinationer av tempus, modus och satstema, vilket exemplifieras i kapitel 8.

De olika teman som anges nedan uttrycks i några fall genom en kod för hela meningen, men i de flesta fall genom ordningskoder, dvs vilken av satserna som logiskt kommer före den andra (1 eller 2). Identifikationen har samband med främst kod 46, men också 42 (nyckelorden är understrukna).

	Kol		Kod
5. Villkor (konditional)	15	Villkorssatsen:	1
		Om jag får fast anställning...	
		Följdsatsen:	2
		så ordnar sig ekonomin	
		Projektet hade avstannat	(2)
		om vi inte fått tillägg	(1)
6. Orsak (kausal)	16	Orsakssatsen:	1
		Eftersom jag hade rätt ålder	
		Följdsatsen:	2
		valde man mig	
		En orsak till att man valde mig	(2)
		var att jag hade rätt ålder	(1)
(jfr kod 46, när fullständig sats inte föreligger)		Det blev mycket jobb	(2)
		till följd av datalagen (46)	(1)
7. Medgivande (koncessiv)	17	Medgivandet:	1
		Det hela avlöpte väl	
		Inskränkningen:	2
		även om det såg illa ut	
		Trots vissa svårigheter (46)	(2)
		kunde arbetet genomföras	(1)
8. Följd/avsikt (konsekutiv/final)	18	Försatsen:	1
		SÖ tog initiativet till det	(1)
		Resultatsatsen:	2
		för att få SIA-underlag	(2)
		Man tryckte på en knapp	(1)
		så att man fick tala med dom	(2)
		För att få veta resultatet	(2)
		gick vi till nämnden	(1)

Anm. Konsekutiva och finala satskonstruktioner har här sammanslagits av den anledningen att det ofta är svårt att i taltexter urskilja de två typerna: talare håller inte isär 'för att' - 'så att'. De finala är dessutom i stor majoritet här. Skulle en uppdelning visa sig nödvändig, kan man använda koderna 3 och 4 och slipper då bryta upp hålkortets disposition.

	Kol		Kod
9. Disjunktion	19	Antingen, visserligen, å ena sidan,...	1
kodas där båda leden förekommer		eller, men, å andra sidan	2
10. Komparation	20	Han (NN) gjorde en bättre sammanfattning	(1)
kodas för förekomst i hela meningen		än vad alla andra lyckades med.	(1)
		Ju mera grundläggande vi gör det (forskar)	(1)
		desto svårare har vi att komma fram med det.	(1)

	Kol		Kod
11. Indirekt fråga kodas för förekomst i satsen. Detsamma gäller direkt fråga (som inte är ställd till intervjuaren), som då suppleras till indirekt konstruktion, men inleds med frågande adverb eller pronomen	21	Jag ville veta om jag kunde resa. Frågan gällde huruvida anslag skulle beviljas. Vi har rader av designer att välja mellan. Vilken av dem är bäst? (Underförstått: undrade vi.)	1 (1) (1)
12. Antagande kodas för förekomst i hela meningen (jfr kod 42)	22	Den (x) skulle förmodligen inneburit besvär. Vi (forskare) hade måst pröva frågan på nytt, antar jag. Jag skulle tro att institutionerna är öppna för flera förslag.	1 (1) (1) (1)
13. Önskan kodas för förekomst i hela meningen	23	Uttrycks på liknande sätt som antagande genom verb eller adverbial (jag hoppas att, förhoppningsvis, önskar, skulle vilja, etc)	1

4.11 Kodning över meningsgränsen

Vi har sett i föregående avsnitt att bisatser är knutna i en logisk följd till huvudsatsen. Denna kodning beror på att meningen är den största analysenheten. Det finns emellertid möjligheter att analysera sådana beroenden utöver meningsgränsen. En sådan typ är redan representerad, nämligen den s k disjunktionen. De två leden utgör varsin huvudsats och kodningen av det första respektive det andra ledet spänner över meningsgränsen (eller -planet), vilket skulle visa sig i identifikationskoderna (gränsen för en ny mening). Det finns även andra möjligheter att gå utanför den formella meningen, nämligen genom att ta fasta på några kriterier, som bortser från huvudsats- respektive bisatskonstruktionen. Denna kodning är då snarare semantisk-funktionell än syntaktisk-formell.

Vi ska här ge några indikatorer för meningsövergripande kodning. Men det betonas särskilt att det bara är fråga om några principer, som inte till fullo har studerats. Vi har i första hand inte velat ta fasta på dessa typer alltför mycket, eftersom det uppmanar till tolkningar, som kan vara svåra att bemästra när det gäller att få tillförlitliga kodningar hos flera personer. Men eftersom det ofta framhålls av forskare som arbetar med liknande problem att "sentence" är den

största enheten som kan kontrolleras, åtminstone när man avser att göra automatisk analys, så bör det här försöket kunna vara vägledande. Vi gör ju ingen automatisk analys på det här stadiet, men varje "nyckel" som vi på manuell väg hittar är naturligtvis till hjälp om man vill vidareutveckla systemet,och vi har, till skillnad från datorn, möjlighet att "förstå" ett yttrande i kontext, vilket bör utnyttjas.

4.11.1 Övergripande teman

För särskiljande av teman inom meningens plan hänvisas till kap 4.10. Nedan exemplifieras endast sådana "nycklar" som inte hänförs till de temastyrande ord som redovisas i kapitel 4.8.6 (kod 46), punkt 3.

1. disjunktion	När inte båda ledens nyckelord finns kan andra ledet ändå innehålla
a) kon(dis)junktioner	eller, men, medan, ...heller, dock, däremot, emellertid, i alla fall, i sin tur, istället, tvärtom, vice versa, å andra sidan, ändå, m fl.
2. orsak	
a) konjunktioner	för, (ty)
b) subjunktioner	därför att (+huvudsatskontruktion)
c) anaforiska adverb	(och) därför (så), därigenom
d) verbkonstruktion	orsaka, vara orsak till, bero på
e) substantivkonstruktion	En orsak (orsaken) till ... var att ...
f) 'med'-satser	med den brist på enhetliga klassificeringar som vi har,...
(jfr kod 46, fraser)	(kan tolkas som på grund av)
3. villkor	
a) verbkonstruktion	med 'skulle' eller vilket annat verb som helst, när satsen implicerar en 'om'-sats: Skulle det vara så ..., (så)... Gör man så,... (så)...
b) substantivkonstruktion	Ett villkor (villkoret)... var att...
4. följd/avsikt	
a) konsekutiva adverb	alltså, följaktligen, följdriktigt, således, sålunda
b) substantivkonstruktion	Avsikten med ... var att ...

5. övrig textbindning

Att en mening har samband med en föregående behöver inte betyda att det finns ett beroende enligt våra punkter i kapitel 4.10. All sammanhängande text knyts ju ihop på ett eller annat sätt. Men följande exempel kan vara tecken på att nå-

got beroende föreligger, som kan kodas i temakolumnerna, dvs en efterföljande mening fortsätter på samma tema.

a) kopulativa och konjunktionella adverb	dels ... dels, desutom, därutöver, förresten, för övrigt, heller, i övrigt, också, likaså, vidare, även, ömsom ... ömsom, m fl
b) konjunktioner	och, men, för, fast
c) understrykande talspråksord	och sen ..., ju, förstås, (se kap 9. 9. 3, punkt 1)

För att uppmärksamma sådana förhållanden, bör man studera främst koderna 42 och 99. En listning kan även göras i efterhand,för att man sedan genom ett register kan koppla dessa nyckelord till vissa textavsnitt,för att se hur de fungerar tillsammans med ett explicit kodat tema (se kap 8).

Anm. Den typ av textbindning som anges via olika slag av bakåtsyftande pronomen behandlas under kapitel 5 om suppleringar.

5. SUPPLERINGSREGLER

I kapitel 3 har vi gått igenom det mest primära ifråga om vårt lingvistiska paradigm, som innebär att vi urskiljer nödvändiga delar för att fullständiga det. Det finns flera sätt att göra yttranden fullständiga och några har antytts vid de olika kodningsreglerna i kapitel 4. I det här kapitlet görs ett försök till sammanfattning av vilka typer av suppleringar som är nödvändiga för vår analys. Typerna indelas enligt de strukturer vi arbetar med (se kap 3). Reglerna är i sin tur beroende av aspekter i materialet, relevansen i förhållande till modellen och analystekniken och möjligheten till teknisk representation för datorbearbetning.

Vi har valt att placera detta kapitel i anslutning till övriga kodningsregler. Men det förutsätts att instuderingen görs i anslutning till exempelsamlingen i kapitel 7.

5.1 AaO-paradigmet i förhållande till satstyp

Man brukar anta tre satstyper, nämligen påstående, fråga och imperativ. Vårt paradigm är av påståendetyp, och de andra typerna kan sägas vara omformade påståendesatser. I kapitel 4.10 redogörs för de olika variationer som satser kan ha. Vi har påståenden med eller utan förbehåll (villkor, orsak, etc.) och vi har särskilda kolumner, som specificerar dessa satstyper och skiljer dem från varandra. För att supplera till fullständiga paradigm (satser) kan vi ange följande grundregler ifråga om satstyp:

1. Påstående

Ett fullständigt påstående-paradigm ska innehålla Agentled, verbled och så många Objektled som verbets betydelse anger. Detta har exemplifierats i kapitel 3 och behöver inte upprepas här.

Det finns emellertid satser, där t ex agentledet inte kan specificeras. Dessa satser kan ändå tas med.

Ex. Biblioteken förbättrades
O a (Agent okänd i materialet.)

Anledningen till att detta ofullständiga paradigm inte utgår är att vi har värden på verben som anger substantivens empiriska vikt.

Anm. Att supplera in de ospecificerade pronomenet 'man' som agent hjälper inte upp paradigmet.

2. Fråga

Vårt material innehåller nästan uteslutande indirekt frågekonstruktion (se kapitel 4.10). En direkt fråga blir inte representerad såsom utsaga i materialet. Den ställs av intervjuaren och det är svaret på frågan som skall kodas.

Ja-nej-svar

Enbart ja, jo, nej eller liknande som svar på fråga måste suppleras med ett helt påstående, jakat eller nekat.

Kodaren kan endast utgå från vad intervjuaren säger och använda den formuleringen.

Ex: Intervjuaren: Har du använt dig av personliga kontaker vid informationssökningen?

Intervjupers: Ja

Suppleringen lyder: Ja (jag har använt mig av personliga kontaker)
resp: Nej (jag har inte...)

(se vidare kod 99, svarsord).

3. Imperativ

Imperativ har fått särskild kod i moduskolumnen (se kap 4.10). Principiellt kan dock anges följande regel:

Imperativen är en satstyp som endast uttrycks via ett verb i imperativform. Agentled saknas. Eftersom imperativ främst används i tilltal till en person suppleras ordet du in (även ni till flera personer).

Ex: 1. du - tilltal	Läs den här boken!
suppleringen lyder:	Läs (du) den här boken
2. ni - tilltal	Jag brukar rekommendera mina lärjungar: Glöm inte dom speciella uppslagsverken!
suppleringen lyder:	Jag brukar rekommendera mina lärjungar (att)* glöm (lärjungar) inte dom speciella uppslagsverken

* 'att' sätts in för att markera att imperativsatsen är ett självständigt paradigm efter suppleringen, samtidigt som 'att'-satsen är objekt till 'rekommendera'. (Se kap 7, exempelsamlingen.) Detta exemplifieras också i kap 5.2.2.

5.2 Relationer

Satsrelationer kan uttryckas genom samordning eller inbäddning (underordning). Som samordnade satser betraktas två huvudsatser, som vanligen är förbundna med 'och'. Två sådana paradigm är oberoende av varandra. Ett inbäddat paradigm är emellertid en beroende enhet till ett annat. Dessa två förhållanden kan uttryckas med begreppen huvudsatser respektive bisatser. I vår kodning av satser och meningar (def. i kap 2) ska dessa förhållanden framträda genom de s k satsslingorna och loopkodningen (se kap 6 och 7).

5.2.1 Supplering vid samordnade AaO-paradigm

1. Om två paradigm har identiska verb, suppleras verbet i det senare paradigmet in och vi får två paradigm (se meningssegmenteringen, kap 2).

Ex: Jag har använt biblioteks sökning och min assistent ERIC-systemet

Suppleringen lyder: (1) Jag har använt bibliotekssökning
(2) Och min assistent (har använt) ERIC-systemet

Konjunktioner och utelämnade verb vägleder kodningen.

Anm. Samordning av Agenter och Objekt vid identiskt verb ger i vår analys inte upphov till två paradigm. Denna samordning anges med ++ (se kap 4.9.1).

5.2.2 Supplering vid inbäddade AaO-paradigm

1. Objekt med infinitiv

a) vid förnimmelseverb (se, höra etc)

Ex: Jag såg honom sitta i biblioteket

Suppleringen lyder: (1) Jag såg
(2) (att) honom sitta i biblioteket
A a

Vad som händer är att den första satsen (1) tar den andra (2) som objekt. 'Honom' (objektsformen) blir Agent och infinitiven 'sitta' är verbet, som får samma tempus som i (1). För att särskilja och ändå relatera de två paradigmen används att som en s k satsinledare (se kap 5.4.2).

b) vid föreställnings- och yttrandeverb (anse, säga etc)

Ex: Lärarkandidaterna ansåg sig vara missbelåtna

Suppleringen lyder: (1) Lärarkandidaterna ansåg
(2) (att) sig vara missbelåtna
A a

(Samma regler för tempus och 'att' som ovan!)

Und. När intervjupersonen själv anser, tycker etc kodas en sådan sats enbart med 42 som fras (se kap 4.8.1 och bil 1).

2. Att-infinitiv innebär att ett nominal (Agent eller Objekt) har uteslutits i den reducerade infinitiv-satsen. Detta nominal återfinns i den överordnade satsen och suppleras.

Ex: Jag lovade honom att skriva en PM

Suppleringen lyder: (1) Jag lovade honom
(2) (att) (jag) (skulle skriva)* skriva en PM

Jfr.: Jag bad honom att skriva en idéskiss

(1) Jag bad honom
(2) (att) (han) (skulle skriva)* skriva en idéskiss

Anm. Verb som 'lova' och 'be' kräver att objektbisatsen får ett 'ska' eller 'skulle', vilket i sin tur beror på tempus (jfr 'Jag ber honom att han ska skriva').

* Parentesens relevans i det här sammanhanget sammanfattas under 5.4.3.

Infinitiven kan också ha en final- ('för att'-) konstruktion.

Ex: Han ringde mig för att få upplysningar

Suppleringen lyder: (1) Han ringde mig
(2) för att (han) (skulle få) få upplysningar

4. Nominalisering innebär att en sats dragits samman till en genitivkonstruktion. Ett sådant nominalkomplex bryts ut till ett fullständigt paradigm.

Ex: X:s ideliga sänkningar av forskningsstandarden oroar stora grupper

Suppleringen lyder: (1) (Att) X:s ideliga sänkningar av Y
A a O
(2) oroar stora grupper

Vad som händer är att nominalet 'sänkningar' bryts ner till verb och att genitivkonstruktionen 'X:s' blir Agent. Det för också med sig att adjektivattributet 'ideliga' förvandlas till adverb, som kodas 45.

Anm. Utbrytning till att-sats förutsätter att genitiven inte uttrycker ägande och att ett fullständigt paradigm uppnås, dvs de tre leden A, a och O. (Jfr kap 4, kod 33 'projektets tillkomst', där agentledet saknas.)

5. Relativsatser

Relativsatserna brukar också kallas attributbisatser. Som attribut tillhör de den efterställda typen (jfr kod 33), och de suppleringar som görs av ett relativt pronomen behandlas i anslutning till supplering inom paradigmet. Den typ av relativsats som anges här avser inbäddning via relativa adverb.

Ex: a) rumsrelatering Jag arbetade inom en grupp där vi hade stor sammanhållning

Suppleringen lyder: (1) Jag arbetade inom en grupp
(2) där (grupp) vi hade stor sammanhållning

b) tidsrelatering Ibland upplever jag perioder då jag inte läser någonting

Suppleringen lyder: (1) Ibland upplever jag perioder
(2) då (perioder) jag inte läser någonting

'Där' och 'då' inleder ny sats och får attributkod som satsinledare och rums- respektive tidsangivelse i den nya satsen.

6. Vilket-satser

Ex: Jag hade jobbat på byrå X i tre år, vilket medförde att jag hade speciella kunskaper

Suppleringen lyder: (1) Jag hade jobbat på byrå X i tre år
(2) vilket (att jag hade jobbat...) medförde
(3) att jag hade speciella kunskaper

'vilket' syftar på hel sats och är satsinledare till denna insupplerade sats, vilken i sin tur är Agent till 'medförde'.

7. Med-fraser

En särskild typ av fraser är sådana som inleds med prepositionen 'med', som tycks sakna innehåll. Det finns där för att tala om att frasen är en reducerad sats. Flera typer kunde anges här, men den typ vi haft anledning att tolka i vårt material är denna:

Ex: Jag har mest begagnat mig av amerikanska tidskrifter med statistisk inriktning

'med'-frasen står i en viss relation till den andra satsen. Vi tolkar den så att den är en bestämning till objektet 'tidskrifter' och kan föras ut i ny sats via ett relativt pronomen 'som'. För att en fullständig sats ska bildas krävs dessutom ett verb, som vi anser vara 'ha'.

Suppleringen lyder: (1) Jag har mest begagnat mig av amerikanska tidskrifter
(2) med (som) (amerikanska) (tidskrifter) (har) statistisk inriktning

5.3 Supplering av strukturer inom AaO-paradigmet

Strukturer inom paradigmet är sådana som återfinns inom Agent-, Objekt- eller verb(adverb)leden. Dessa strukturer avspeglar sig i kodsiffrorna (kap 4) och angavs principiellt i kapitel 2. Suppleringar här avser i första hand huvudbegreppen (kodsiffra 0), men även bestämningar kan komma ifråga.

5.3.1 Pronomen

Pronomen suppleras med sina korrelat, dvs de begrepp som är deras referens i texten, när denna referens inte anges i samma sats.

Ersätter huvudbegrepp

1. personliga	jag, min, mig, du, din, dig, ni, er, han, hans, sin, honom, hon, hennes, henne, den, (dess, dens), det, vi, vår, oss, ni, er, de (dom), deras, sin, sig, dom (dem), man
2. demonstrativa	den där, den här, etc
3. relativa	som, vad
4. reciproka	varann, varandra

Kan ibland ersätta huvudbegrepp

5. totala	alla, båda, inga, ingenting, etc.
6. kvantitativa	många, mycket, några, någonting, några få, etc.
7. selektiva	övriga, den senare, etc.
8. komparativa	sådan, likadan, dylik, etc.
9. limitativa	somliga, vissa, andra, en del, en annan, några, etc.

Jag (min, mig) har i vårt material inte supplerats, eftersom identifikationskoden säger vilken person det gäller. En rutin har införts senare, som maskinellt specificerar jag-referenser med ip-nummer.

I en talsituation av det här slaget förekommer det ofta att intervjupersonen säger man, vi eller du istället för jag. Om det är klart utläsbart suppleras (jag) i dessa fall.

Vi är i vårt material ofta liktydigt med 'projektmedlemmar', 'projekt' eller 'institution'. I sådana fall är det lämpligt att med utgångspunkt i materialet specificera med lämpligt begrepp. I vårt fall har vi dessutom tillgång till projektnamn med förkortningar, varför vi supplerats med t ex 'xyz-projektet'. (För att underlätta för kodarna medföljde en lista över projektnamn, institutionsförkortningar, etc.)

När någon, annan etc förekommer utan klar referens suppleras det med 'person'. (Förväxling med någon + 'sak' ska inte ske.)

Vissa pronomen ersätter adjektiv (t ex sådan, likadan) (jfr. kap 4.2.2, punkt 5).

Anm. Reflexiva pronomen (former med 'sig') går i vår analys in i verbet (se kap 4.1, punkt 46).

5.3.2 Proformer - betydelsefattiga ord

1. verbet 'göra'

Ex:	Vi har gjort en avgränsning av problemområdet
Suppleringen lyder: (1)	Vi har gjort en avgränsning (avgränsat) av problemområdet.

Vårt material innehåller beskrivningar av olika verksamheter, varför former av 'göra' supplerats med 'arbeta med', 'undersöka', etc. Objektet (med suffixet -ning) blir verb och det begrepp som inleds med prepositionen 'av' blir Objekt.

Anm. Om 'göra'= 'tillverka' behövs ingen supplering.

2. verbet 'vara'

'vara' som kopula har kod 41. Om 'vara'= 'befinna sig' och en rumsbestämning följer, behöver det inte suppleras, eftersom det kodas med 40. Det finns dock

anledning att vara uppmärksam på om 'vara' betyder befintlighet eller om det kan suppleras med ett mera aktivt verb (se punkt 3 nedan).

3. Verbs betydelse

Sammanhanget avgör om verbet kan innebära en tvetydighet. I vissa fall suppleras verbet med ett lämpligare.

Ex:	En projektbeskrivning gick in (skickades) till SÖ (av xyz-projektet)
Några exempel i intervjumaterialet:	Jag gick på (sökte på) Abstracts Det har blivit (uppstått) svårigheter Han kom med (utgav) en bok Jag vet inte vad ERIC är (känner till) Vi har varit på (besökt/bevistat) symposier Dom visste inte vad vi gjorde (sysslade med) Jag har fått igenom (presenterat) högar med rapporter

Vi har inte närmare angivit regler för vilka verb som ska suppleras med vilka, i första hand därför att antalet inte varit särskilt stort. Exemplen ovan kan vara vägledande och är tänkta att göra kodaren uppmärksam på typen av uttryck.

4. Substantiv

Här kan endast en allmän regel ges: Undvik sådana ord som, när de ses isolerade i en substantivlista, inte är specifika.

Några exempel: (se även kod 46)	sak, grej, typ, område, etc

Suppleringen tar hänsyn till det ämnesområde som texten behandlar.

Ex:	Vi har haft hand om många saker
Suppleringen lyder: (1)	Vi (xyz-projektet) har haft hand om många saker (undersökningar)
	Det här området har inte studerats så ingående
Suppleringen lyder: (1)	Det här området (tvåspråkighet) har inte studerats så ingående

5.3.3 Pronomen som ersätter hel sats

1. Det

Det förekommer ofta att intervjupersonen svarar med ett 'det' som innefattar en hel sats och syftar på den fråga som intervjuaren ställer.

Ex: Intervjuaren:	När började du bli intresserad av det här problemområdet
Intervjupersonen:	Det började när jag läste i Lund.

Här måste vi veta dels vilket problemområde den här personen arbetar med, dels hur suppleringen av 'det' ska ske. Eftersom materialet är omfattande och har ett stort antal 'det'-referenser av det här slaget, upprättade vi en lista på standardiserade suppleringar, dvs hur suppleringen skulle ske till varje fråga. Denna lista var en del av kodningsinstruktionerna. Vi tog fasta på 'huvudsaken' i frågan och skapade en struktur med huvudbegrepp + bestämning. Lösningen på exemplet ovan blir:

Det (intresset) (för barns uppväxtmiljö)
började när jag läste i Lund.

Genom standardiseringen undviks en alltför olikartad kodning, vilket är en risk vid kodning genom flera personer. Det är dessutom arbets- och utrymmesbesparande att slippa göra hela satser i alla dessa fall som återkommer flera gånger hos en och samma person.

5.3.4 Relativsatser

I det här sammanhanget betraktas relativkonstruktioner som en del av paradigmet, eftersom de är attributiva (se kap 4.2.3). Vi skiljer inte på restriktiva och icke-restriktiva, eftersom vi behandlar talat språk i nedskrivet skick, där denna skillnad inte så lätt framträder. (Se ex i kap 3.3.) Relativa adverb behandlas under 5.2.2.

1. som (vilken)

Ex: Rapporten som kom igår har jag inte hunnit läsa

Suppleringen lyder:
(1) Rapporten
(2) som (rapporten) kom igår
(3) har jag inte hunnit läsa

Ex: Ett projekt som kallades Skrivsyntax startade

Suppleringen lyder:
(1) Ett projekt
(2) som (ett projekt) kallades Skrivsyntax
(3) startade

Satserna är följande:
Rapporten har jag inte hunnit läsa
Rapporten kom igår

respektive:
Ett projekt startade
Ett projekt kallades Skrivsyntax

Den andra satsen i respektive exempel bäddas in i den första och kodas som attribut (efterställt), där 'som' är inledare till den inbäddade satsen. (Exempel på detta finns i kap 7.)

2. vilket — syftar på hel sats, se 5.2.2, punkt 6.

3. vad

Ex: Jag utsattes för hård kritik utan att jag visste vad som var fel

Suppleringen lyder:	(1)	Jag utsattes för hård kritik (av X)
	(2)	utan att jag visste vad (fel)
(se kodn., sid 73)	(3)	som var fel

5.4 Tekniska aspekter vid supplering

5.4.1 En teknisk kod

Vid fall av flätning, där objektet i den inbäddade satsen står först, t ex vid emfas, tillgrips en teknisk kod som möjliggör en utbrytning i ny sats.

Ex: Pedagogiken ansåg dom vara intressant

Suppleringen lyder:	(1)	Pedagogiken ansåg dom 88
	(2)	(att) pedagogiken vara intressant 50

Intervjupersonens egen mening (jag anser, etc) föranleder ingen utbrytning:

Ex: Pedagogiken ansåg jag vara intressant
30 42 41 32

Fall av flätning exemplifieras närmare i kapitel 7.

5.4.2 Satsinledare

För att kunna markera en ny sats inom en mening har vi ett s k looping-system, där vi anger i vissa kolumner och med vissa siffror att ny sats börjar. Looping-systemet exemplifieras i kapitel 6.2.2. Här ska vi endast repetera vilka typer som är satsinledare (gäller inte samordning).

1.	att (hur, om+fråge-ord)	(har ej eget innehåll)	Inleder Agent- eller Objektsats. Hela satsen tillhör paradigmet och får kodsiffra som slutar på 0.
2.	som vad	(har eget innehåll)	Inleder attributsats. Hela satsen tillhör paradigmet och får kodsiffra som slutar på 3. I den nya satsen kan första siffran vara en annan än i föregående sats.
3.	tids- och rumsadverb (relativa)	(har eget innehåll)	Inleder tids- och rumssatser. Hela satsen tillhör bestämningen (43 resp 44).
4.	ord eller fraser som anger		
	a) sätt/grad	(har ej eget innehåll)	Inleder sätt/grad-adverbialssatser Hela satsen tillhör bestämningen (45).
	b) satstema	(har ej eget innehåll)	Inleder satser som anger komparation, orsak, villkor, följd, avsikt, medgivande, disjunktion (46).

5. vilket (har ej eget innehåll)	Inleder ny sats genom att syfta på en hel sats. Som satsinledare får ordet kod 46, följt av att, som kodas med innehåll.

5.4.3 Parentesernas innehåll

Anvisningar för uppskrivning på datablad görs i kapitel 6.3.1. Här ska vi nämna några principer som är resultat av vårt arbete med lexikon.

1. Parenteser används för att supplera sådana enheter som, enligt reglerna i detta kapitel, är relevanta inom och mellan paradigm.
2. Inom parentes står den huvudsakliga enheten för varje hålkort. Det betyder att

   a) de ord som står utanför parentesen betraktas som irrelevanta, och att

   b) tillägg till ett ord genom parentes inuti ordet inte kan göras, samt att

   c) hjälpverb inte får kodas inom parentes på annat kort än huvudverbet

   Ex: "(ska kunna) läsa" ger inte utfall på 'läsa'.

Om hjälpverben också har betydelse måste hela frasen stå inom parentes: "(ska kunna läsa) läsa" (Jfr 5.2.2, punkt 3). Vårt datorprogram har instruktionen att det sista ordet inom en parentes är det som koden gäller. Detta har inneburit att vi kunnat formulera en tredje princip

3. Tempusbetecknande hjälpverb suppleras inte. Om sådana finns i texten, har det ingen betydelse, eftersom inga parenteser komplicerar. (Tempusbeteckning finns i kolumn 13.)

Det hjälper också till att tolka verbet 'ha' i betydelsen 'äga'. Ett ensamt 'har' på ett kort utan parentes tolkas som 'äga', om verbkoden inte är delad. Förekommer ett huvudverb i satsen, gäller detta verb.

5.5 Sammanfattande nyckel

5.5.1 Suppleringar i förhållande till materialet

1. Svar till intervjufrågan	Svarsorden får kod 99 och svaret suppleras till fullständigt paradigm. Intervjuarens ordval väljs.
2. Intervjupersonens åsikter	'Jag tycker att'... och liknande fraser får kod 42. Endast referenser till någon annans åsikt ('Han (X) ansåg att'...) kodas som eget paradigm.
3. Ordvalet	Suppleringar av pronomen, verb och pro-former ska ansluta sig till kontexten och ämnesområdet. Tvetydigheter undviks i möjligaste mån. Standardiserade begrepp kan bestämmas om frekvensen berättigar till det.

5.5.2 Suppleringar i förhållande till modellen

1. AaO-paradigmet	Supplera till fullständighet genom att kontrollera verbets betydelse och funktion.

2. Förhållandet mellan paradigm	Varje paradigm ska ha ett verb. Vid flera verb ska relationerna mellan paradigmen suppleras via satsinledare. Har ett verb uteslutits kan det vara fråga om samordning. Verbet suppleras då in och satserna utgör två meningar.
3. Satsreduktioner	Kontrollera om verbet är av typen förnimmelse, föreställnings eller yttrande, eller står i infinitiv. Kontrollera också om nominalisering föreligger. Supplera till inbäddade paradigm.
4. Enskilda begrepp	Kontrollera varje pronomens eller adverbs referens. Supplera om kontexten är tolkbar:
a) i första hand	huvudbegrepp, slutsiffra 0
b) i andra hand	beskrivningar, slutsiffra 2
c) i tredje hand	bestämningar, slutsiffra 3
Suppleringens omfång	Om ett pronomen syftar tillbaka på ett helt komplex (t ex 51, 52, 50) i föregående sats, suppleras hela komplexet in i nästa sats. Det räcker inte med huvudordet.
Suppleringens form*	Välj samma form (bestämd, obestämd) som referenten. Själva artikeln kan utelämnas.
Indikationen för supplering (allmänt)*	Supplering sker oftare vid bestämda former än obestämda.

* Det är inte motiverat att här gå in i detalj på frågan om "diskursreferenter" (känt och okänt för den som lyssnar). Våra lexikon utgörs av former utan ändelser och man kan välja att supplera i obestämd form (grundform) likaväl som bestämd form. Men när substantiv används i bestämd form (känt för lyssnaren), kan det finnas anledning att kontrollera substantivet ifråga om dess specificerade betydelse (jfr pro-formerna), t ex att ett tillägg (kod 33) kan behöva göras.

5.5.3 <u>Suppleringar i förhållande till tekniken</u>

1. Teknisk kodning (satskolumner)		Kontrollera innebörden i satsinledarna, samt flätningar.
2. Parentesers innehåll		Bestäm vad som är det huvudsakliga innehållet i den enhet, där supplering sker, så att oväsentliga delar går utanför parentesen. Vid verbfraser - kontrollera hjälpverbens sammansättning. Sista ordet i parentesen gäller.
3. Avgränsning av begrepp		En parentes får inte innesluta två begrepp med olika koder, som, om de funnits utanför parentes, skulle ha uppdelats på olika kort. Detta avgörs dock i förhållande till analysens syfte.
I vårt fall gäller:	a)	Substantiv hålls isär från sina bestämningar, om substantiven är A eller något O.
	b)	Efterställd bestämning, som består av flera substantiv, avgränsas genom att de kodas på varsitt kort.
	c)	När ett bestämningsord som skulle få slutsiffra 1 eller 2 ska suppleras, sätts det supplerade huvudordet på nytt kort.
<u>Undantag</u>		Pro-substantiv (och verb) suppleras på samma kort som det ord som de ersätter.

6. DATABLAD

6.1 Disposition

För utskrift av text tillsammans med kodmarkeringar används en blankett i A4-format (datablad), försedd med 80 kolumner, dvs samma antal som hålkortet, på vilket materialet stansas. Varje kolumn kan markeras antingen med bokstavstecken eller likvärdigt eller också med siffror från 0 till 9. Kolumn som i vissa fall lämnas blank har också en kod.

En god princip när det gäller uppdelningen av olika slag av koder är att skilja på numeriska och alfabetiska, vilket man också internationellt har försökt göra till en konvention vid detta slag av analyser. De koder som är viktigast för att sortera texten i större avsnitt bör finnas i början, eftersom det då tar mindre tid för datorn att utföra instruktionerna.

Längst till vänster på bladet skrivs identifikationskoder (kol 1 - 11), därefter följer koder för olika satsteman (kol 12 - 23). Kolumnerna 24 - 25 används när ett ord löper över flera kort än ett och har samma kodbeteckning. (Detta förfarande beskrivs utförligare nedan under 6.2). Kolumnerna 26 - 65 upptar texten och de sista (66 - 80) anger de syntaktisk-funktionella koderna (se fig 1).

6.2 Tekniska anvisningar utom textkolumnerna

Det är både tidsbesparande och överskådligt för kodaren om identifikations- och temakoder skrivs upp enligt en förenklad princip (se fig 2). Det har visat sig vid våra utprövningar att följande tillvägagångssätt fungerat bra:

1. Kolumnerna 1 - 8 skrivs upp en gång vid varje ny menings början.
2. Kolumnerna 9 - 10 används så att nollan i ordnummer 01 anges första gången och inte upprepas. När man kommer till nummer 10 fylls givetvis båda kolumnerna. Kolumn 11 anges som temat.
3. Temakolumnerna anges med respektive kod vid menings början eller på det ställe i meningen där ny kod skall anges, om den inte är den samma som i början. Så snart ny kod följer markeras denna på den rad där den nya satsen börjar och då gäller den senare koden tills den bryts. Koden bryts alltid vid ny mening, även om den skulle vara densamma.

   Vi använde dessutom en princip vid negationen (kol 12) nämligen att ange ett streck (-), som betyder 'blank'. Detta gjordes för att markera att kodningen inte hoppats över. Negationer är viktiga.

   OBS! Raden på bladet fylls nedåt med den kod som angivits vid menings eller sats början i stansningen! Alla hålkort ska alltså ha dessa koder, inte bara det första i varje mening. (Figur 3 visar hur hålkorten ser ut efter stansning av texten i figur 2.) Detta ska noggrant anges

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard: EBCDIC, BCD | Speciell stans-instruktion | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar

FORTRAN: Fml nr 5 | 6 | 7 FORTRAN-formel | Identifikation / Sekvens

DATA | ALGOL Sats | PL/I Sats

1–11	12–25	26–65	66–80
3808d032011	1-11	Blanketten	30
2		visar	40
3		ett enkelt	52
4		kodningsexempel	50
5		som (kodningsexempel)	53230
6		anger	40
7		disposition	50
8		och	++
9		förenkling	50

Identifikation

1-2 intervjupers. nr
3-4 fråga nr
5-8 mening nr
9-10 enhet nr i men.
11 källa t. utsaga

Tema *

12-14 negation, tempus resp modus (oblig. kodn.)
15-23 olika teman kodas vid förekomst

Funktionella enheter (ord)

26-65 textlängd per enhet
24-25 * används när text löper över flera kort än ett

Syntaktisk funktion

66-80 innehållsl. funktion
68, 71, 74, 77, 80 strukturell funktion (loop)

Figur 2. Datablankett: Disposition och kodning

M7102 520070. Utgåva 1: 1970 april. 1000 bl à 100 ex (FOA 404)

```
38080032011 11    BLANKETTEN              30
38080032021 11    VISAR                   40
38080032031 11    ETT ENKELT              52
38080032041 11    KODNINGSEXEMPEL         50
38080032051 11    SOM(KODNINGSEXEMPEL)    53230
38080032061 11    ANGER                   40
38080032071 11    DISPOSITION             50
38080032081 11    OCH                     ++
38080032091 11    FöRENKLING              50
```

Lunds Datacentral

UNIVAC 1108

Figur 3. Hålkort: Stansad text

för den som ska utföra stansningen. Specificera i anvisningarna vilka kolumner detta "ekonomiska" kodningsförfarande gäller eller andra systematiska principer. (Exempel på stansinstruktion ges i bil. 2:2.)

Dels underlättar detta förfaringssätt betydligt när man sitter och arbetar med en text, eftersom de flesta av dessa koder anges ganska mekaniskt och blir tröttande, dels har man en god överblick över databladen vid kontroller, eftersom man lätt ser på vilka ställen ett eventuellt fel har gjorts. Visserligen kan flera av dessa kolumner ifyllas av icke skolad personal, men det är ett sämre alternativ och ökar risken för felskrivning, som kan bli besvärligt att kontrollera.

6.2.1 Kolumnerna 24 - 25

I de fall ett begrepp eller följd av ord inte får plats på ett kort och har samma syntaktiskt-funktionella kod, används två kort (rader), som av datorn ska betraktas som samma kort. I kolumnerna för ordnummer (radnummer) skrivs då samma kod för begreppet. Dessutom anges i kolumnerna 24 och 25 ordenheternas inbördes ordning enligt följande.

Kolumn		eller	kolumn	
24	25		24	25
2	1		3	1
2	2		3	2
			3	3

Kolumn 24 anger hur många kort som hör ihop och kolumn 25 vilken ordning korten har av två respektive tre. Kolumnerna lämnas tomma då de inte används.

OBS! Avstavning med bindestreck mellan två kort får inte förekomma (se nedan under 6.3)

6.2.2 Kolumnerna 66 - 80

Satskolumnerna innehåller de syntaktisk-funktionella koderna (se kap 4). Varje sats börjar genom att en siffra hänvisar till den nya satsen (loop). Datorn måste få anvisning om var den ska leta efter fortsatt text, om texten "tar slut" utan att meningen är slut. Siffran i loop-kolumnen kan visa bakåt eller framåt. Det kan nämligen inträffa att ett ord i den första satsen först har en bestämning i form av bisats i ny kolumn och sedan fortsätter i den första kolumnen. Den första satsen (kol 66 - 67) är huvudsatsen, men oavsett i vilken ordning satserna följer får varje satskolumn inte innehålla mer än ett huvudverb.

Genom detta kodningsförfarande har kodaren överblick över hur satserna löper i slingor, och denna visuella bild över texten i maskinutskrivet skick (kortbilden) underlättar betydligt vid kontroller. Genom att samtidigt se satsslingorna och identifikations- och temakoder går det lätt att uppfatta om felkodningar finns.

6.3 Tekniska anvisningar inom textkolumnerna

För detaljer angående de syntaktisk-funktionella koderna och identifieringen av de olika begreppen inom en mening hänvisas till kapitel 4 och 5. Detta avsnitt behandlar endast sådana aspekter som inte har att göra med innehållsliga överväganden i texten. I princip skrivs varje enhet på en rad. När det gäller avgränsningen av enheter ska emellertid följande påpekas:

1. Om två på varandra följande enheter var för sig kan bilda en oberoende enhet, även om de har samma kod, skrivs de på varsin rad.

   Ex: Han (NN)
   satt
   på en gammal stol ↖ två enheter
   på sitt arbetsrum ↙

2. Uppräkningar som kan skiljas åt genom 'och', 'eller' eller kommatecken (i skrift) skrivs på varsin rad, även om de har samma kod.

   Ex: Jag
   sysslade med
   lågstadielärare ↖
   eller ↘ två enheter
   mellanstadielärare ↙

3. En enhet kan brytas genom en annan och sedan fortsätta.

   Ex: började ←
   jag
   med → hör till verbet
   projektskissen

När enheterna skrivs upp på databladet gäller följande regler:

4. Texten skrivs med en bokstav i varje kolumn och med tom kolumn efter varje ordslut.
5. Jämförbara med bokstäver är bindestreck, accenter och parenteser. Accent skrivs i egen kolumn, t ex ide'. (Bindestreck och parentes behandlas nedan under 6.3.1.)
6. Siffror skrivs med bokstäver, t ex 'två böcker', utom årtal.
7. Förkortningar kan användas, men man bör avgöra om de är tvetydiga eller ej. Datorn har lättare att behandla förkortningar om de sammanskrivs, t ex 'osv', 'dvs'. Skulle en förkortning av en händelse bli liktydig med ett annat ord, skrivs den med mellanrum eller enligt någon annan princip som man kommer överens om.

8. Förkortningar på t ex institutioner eller organisationer kan användas, såsom LHM, SACO, SÖ.
9. Interpunktion används inte. Sådana funktioner anges genom identifikations- eller temakoder. (Interpunktionstecken är bara av betydelse om man analyserar skriven text utan att segmentera i enheter.)

Som en anmärkning till dessa punkter ska sägas att förutsättningarna och därmed reglerna kan vara annorlunda om stansningen sker med mera tekniskt avancerade resurser. Det finns t ex numera maskiner som stansar stora och små bokstäver eller har flera teckenkombinationer än vad vi här utgått ifrån. Reglerna bestäms utifrån den utrustning man har tillgång till.

6.3.1 Bindestreck och parentes

Det har visat sig vid våra olika utprövningar på det stansade materialet att bindestreck och parenteser måste användas med eftertanke. För bindestreck gäller följande:

1. En sammanlänkning av två ord genom bindestreck används inte, t ex i typen 'låg-och mellanstadielärare'. Begreppen delas upp i två fullständiga begrepp 'lågstadielärare' resp 'mellanstadielärare'. Uppdelningen innebär därmed att de två begreppen skrivs på varsin rad och samma regler som för samordning gäller (se kap 4.9).
2. Bindestreck som har till funktion att binda ihop två begrepp till ett får förekomma, om avsikten är att betrakta detta ord som ett istället för två, t ex 'mano-depressiv'.
3. Vid uppdelning av begrepp enligt principerna för kolumnerna 24 - 25 (se 6.2.1 ovan) får bindestreck inte vara avdelare av ett begrepp (avstavning). Inte heller får bindestreck enligt punkt två förekomma.

   Principen för avstavning gäller ju när ett ord fortsätter och när ordet inuti en textrad skulle sammanskrivas. En sådan avstavning kan inte hanteras av datorn. När datorn läser ihop två kort inträffar det nämligen att blank kolumn följer på bindestrecket och bokstavsföljden bryts, vilket innebär att en sökning på begreppet ifråga inte ger utfall.

Parenteser har i våra regler en särskild betydelse, nämligen att ange supplering (tillägg) av ord enligt vissa kriterier (se kap 5).

Principiellt gäller:

1. Något mellanrum mellan parentes och övrig text görs inte. Texten innanför parentesen börjar utan mellanslag.
2. Om uppdelning av begrepp i flera kort sker och ett begrepp med parentes ingår, kan parentesen inte börja på ett kort och sluta på nästa. Hela begreppet inom parentes skall föras till ett kort. Detta gäller oavsett om begreppet har bindestreck eller ej.

6.4 Förberedelse för stansning

När uppskrivning av text tillsammans med koder har skett och eventuellt kontrollerats ska materialet stansas på hålkort eller hålremsa, eller vilka andra databärande media man har tillgång till; det beror ofta på mängden och därmed lagringsmöjligheterna. Vårt material utgår från hålkort. Andra media kan medföra att den disposition (s k "format") vi utformat får anpassas på lämpligt sätt.

Vi har gjort den erfarenheten att noggrant förberedda anvisningar till stansoperatrisen är av största betydelse. Otydligt skrivna datablad eller kodningsscheman innebär större arbete med rättningar efteråt. Stansresultatet blir också bättre om operatrisen informeras om vilken typ av text det gäller och vad koderna i huvudsak innebär, vilket gör det lättare för honom eller henne att själv upptäcka fel, åtminstone av enklare slag, t ex stavfel. Men viktigast är att ange vilka slags koder som kan förekomma inom vilka kolumner på blanketten (hålkortets disposition). Stansoperatrisen bör ha kodningsschema och stansinstruktioner. Detta exemplifieras för vårt material i bilaga 2.

7. EXEMPELSAMLING MED KOMMENTARER

I detta kapitel exemplifieras reglerna genom kodning på datablad av meningar, satser, begrepp och tekniska lösningar. Exemplen är till största delen hämtade från vårt intervjumaterial. Identifikationen och vissa suppleringar är emellertid fingerade.

Samlingen avslutas med kodning av den text vi presenterade i ruta 1. Den är inte kommenterad utan bör kunna förstås utifrån reglerna och kodningsexemplen på de föregående bladen.

Vi kan inte vara alltför omfattande i den här exempelsamlingen. Alla varianter och problemtyper blir inte representerade. Vi får istället hoppas att läsaren (kodaren) kan utnyttja den generalitet som kan återspeglas i regeluppsättningen. I mångt och mycket skulle nog kodningen kunna förenklas och det står var och en fritt att upptäcka sådana fall, liksom också att bidra med egna lösningar till sådana empiriska fall som inte kommit med här. Författaren till den här versionen av manualen tar tacksamt emot tips för eventuella kommande revideringar.

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard: EBCDIC / BCD | Speciell stans-instruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRAN-formel
DATA / ALGOL Sats / PL/I Sats
Identifikation / Sekvens

Kol. 1–24	Kol. 26– (text)	Kol. 65–71	Kol. 73–
0101 0010 11-25	Jag	30	
2	började jobba med	40	
3	det här (invandrarproblem)	50	
4	på hösten 1970	43	
201 1-21	Men	99	
2	jag	30	
3	hade intresserat mig för	40	
4	sina frågor (invandrarproblem)	50	
5	långt tidigare	43	
6	under min studietid	43	
7	i USA	44	
301 1-21	När	432	
2	jag	30	
3	kom tillbaka	40	
4	till Sverige	601	
5	fick	40	
6	jag	70	
7	en assistenttjänst	50	
8	i det här projektet (JP-projektet)	44	
9 3 1	som (assistenttjänst)	533	50
10	jag		30
11	har (innehar)		40
12	i två år till		43
13	förmodligen		42

Kommentarer i marginalen:

- hjälpverb + huvudverb
- supplering av problemområde
- verbfras
- ny mening
- två likvärdiga begrepp
- satsinledaren visar direkt till andra kol
- temporal bisats
- Objekt(Mot) vid "få" (agentlöst)
- hänvisning till huvudsatsen
- tempus: framtid
- supplering vid relativt pron
- "som" har egen betydelse
- supplering vid tvetydigt verb
- antagande
- hänvisning till närmast lediga satskolumn

M7102-62000 3. Utgåva 1: 1976 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kyndnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Speciell stansinstruktion | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar | Identifikation / Sekvensnr

FORTRAN: Fml n | FORTRANformel

DATA / ALGOL / PL/I

ny fråga ↓

Kol. 1–10	Kol. 11–20	Kol. 25– (Operand)	Kol. 66–70
0105000101	1-11	Jag tycker att	42
2		det här (informationssökning)	30
3		är	41
4		väldigt problematiskt	32
201	1-21	Nog	42
2		har	40
3		jag	30
4		väl	42
5		använt mig av	40
6		ERIC-basen	50
7		en del	43
301	1-11	Men	99
2		å andra sidan	46
3		innehåller	40
4		den (ERIC-basen)	30
5		mängder av	51
6		information	50
7	1 4	som (information)	53 2 50
8		man (jag)	30
9		inte	--
10		har möjlighet att läsa	40
401	1-11 1	Psychological Abstracts	30
2		ger	40
3		mig	70
4		bättre	52

Anteckningar på blanketten:

- satsadverbialssats (pil till "att" och till koden 42)
- adverb + adjektiv i samma kod (pil till "väldigt problematiskt")
- frekvens (pilar till "en del")
- obs! koden ges utan temakod för disjunktion (pil till koden 46)
- negation (pil till kolumn 12 i rad 7)
- modal betydelse tolkning = "kan" (pil till kolumn 14 i rad 7 och till "har möjlighet att läsa")
- "jag"-referens (pil till "man (jag)")
- riktning! vem agerar (pilar till koderna 30 och 70)
- komparation (pil till kolumn 20 i rad 401)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard | EBCDIC | BCD | Speciell stans-instruktion | Tecken | Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar | Identifikation / Sekvens

FORTRAN: Fml nr 5 | 6 | 7

DATA ALGOL: Sats

PL/I: Sats

ifyllning vid nytt datablad

Kolumn 1–10	11–15	20	25– (Operand)	66–
0105000405	1-11	1	överblick	50
6			(över litteratur)	53
7			(än ERIC-basen)	46
501	1-21		Annars	99
2			kan jag väl säga att	42
3			jag	30
4			sökt	40
5			(litteratur)	50
6			med hjälp av flera	81
7			olika	81
8			referatorgan	80
9			som(referatorgan)	83 2 30
10			har varit	41
11			mer eller mindre bra	32
601	1-21		Men	99
2			(jag)	30
3			(har sökt)	40
4			(litteratur)	50
5			oftast	43
6			manuellt	45
701	111		Jag	30
2			tror	40
3			inte	--
4			på	40
5			datasökningssystemen	50

Anmärkningar:

- (än ERIC-basen) ← supplering av komparationen
- 46 ← komparation
- med hjälp av flera / referatorgan ← Objekt(Med): sista ordet gäller för koden. Frasen är "nyckel"
- 81 / 81 ← två likvärdiga begrepp
- 83 2 30 ← annan betydelse i ny sats
- 41 / 32 ← predikativ
- (jag) / (har sökt) / (litteratur) ← supplering av paradigmet
- tror ← verbet "tro på" ej satsadverb.

M7102-620070. Utgåva 1. 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Speciell stansinstruktion | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar | Identifikation / Sekvensnr
FORTRAN: Fml nr 5 | 6 | 7 FORTRAN-formel
DATA
ALGOL Sats
PL/I Sats

1–10	11–15	25–	63–70
0105000706	1111	eller	++
7		övriga	51
8		mekaniserade	52
9		rutiner	50
0202000101	1-11	Avgränsningen	30
2		(inom XYZ-projektet)	33
3		innebär	40
4		i stort sett	42
5		att	502
6		jag	30
7		skriver om	40
8		intervjumaterialet	50
9		när det gäller skolledarbiten	53
10	3	som(skolledarbiten)	533 30
11		leder fram till	40
12		en beskrivning	50
13		av vad typ av arbetsuppgifter	53
14	1	som(arbetsuppgifter)	534 50
15		skolledare	70
16		ställs inför	40
201	2-21	Dom(skolledarna)	30
2		menade	40
3		att	502
4		uppgifterna	30

Annotations:

- konjunktivt "eller": bindning av objekt
- ny ip
- att-sats: "att" har ej egen betydelse
- ändring av kod vid satsinledaren
- tempus-brytning
- två objekt vid verb på -s
- annan källa än ip vid "mena", etc = ej satsadv

A7102-620070. Utgåva 1 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Speciell stansinstruktion | Tecken | Stansning | Sida ... av ... | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel
DATA
ALGOL Sats
PL/I Sats

Identifikation / Sekvens (73–80)

obs! kolumn följs

Kol. 1–10	Kol. 11–15	Kol. 26–	Kol. 68–72	Anmärkning
0201000205	2-21	skiftade	40	
0301000101	1-11	Det (att)	302	
2		(projektstarten)	30	
3		(ligger)	40	
4		(en bit tillbaka i tiden)	431	supplering av hel att-sats; att-satsen är bisats
5		innebär	40	
6		glömska	50	
7		och	++	
8		rationaliseringar	50	
201	1121	Men	99	
2		såvitt jag kan komma ihåg	42	
3		fanns	40	
4		det	99	innehållslöst "det"
5		inget (forskning)	50	negation; suppleras på samma rad
6		gjort	40	tolkning = hade gjorts
7		där (inom YB)	44	
8		då (vid projektstarten)	43	
301	1-21	Ett hundratal	31	
2		experter	30	
3		satt (forskade)	40	
4		väl	42	
5		i sju år	43	
6		(i YB)	44	

AV102-620070. Utgåva 1. 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Specialstansinstruktion | Sida av | Program

Assembler (angi Namn) / FORTRAN / DATA / ALGOL Sats / PL/I Sats | Operand | Kommentar | Identifikation / Sekvensnr

Id (kol. 1–9)	Rad	Tempus	Avsikt/följd	Önskan	Operand	Kod
040700010	1	1-21		1	Jag	30
	2				hade	40
	3				länge	43
	4				önskat	40
	5	4			att	502
	6				litteratursökning	30
	7				kunde ske	40
	8		1		på det sättet att	453
	9				man (forskare)	70
	10	3			finge	40
	11				välfungerande	52
	12				uppslagsord	50
	13				(från referatorganen)	30
	14	4			där (i uppslagsord)	53444
	15				man (forskare)	30
	16				skulle kunna se	40
	17				någon struktur	50
	18	5	2		för att	465
	19				(forskaren)	30
	20				gå vidare på (fortsätta söka)	40
	21				den (i strukturen)	44

Anteckningar:

- tempus för huvudsatsen gäller även vid de övriga modus
- avsikt/följd
- önskan
- konjunktiv form
- modalt hjälpverb
- avsikt/följd - gäller ej hela men.
- ny sats - supplering av Agent utan pronomenet
- inleder sätts-sats
- supplering med material-typisk referent
- Agenten känd
- rel pron inleder och byter betydelse
- obest artikel på samma rad
- satsinledare för tema

M7102-620070. Utgåva 1: 1970 april. 1000 bl à 100 ex (POA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Speciell stansinstruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar

FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel

DATA ALGOL Sats PL/I Sats 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 71 | 73 Identifikation / Sekvensor | 80

1–9	10	11–14	16–20	Operand (25–)	65–73
040700020	1	1-11	11	Visserligen	46
	2			är	41
	3			det	99
	4			bra	32
	5	4		att	302
	6			man (forskare)	30
	7			kan gå ner	40
	8			till biblioteket	60
	9		2	och (för att)	463
	10			(forskare)	70
	11			få hjälp	40
	12			med litteratur	50
	13			(av biblioteket)	30
30	1	1-11	2	Men	99
	2			å andra sidan	46
	3			är	41
	4			det	99
	5			ju	42
	6			fint	32
	7	4		att	302
	8			vi (forskare)	30
	9			kan sitta	40
	10			på rummet (arbetsrummet)	44
	11			och läsa	40
	12			bibliotekets	51

disjunktion – två men.

att-satsen värderas

satsförkortn. som görs till avsiktssats

verbet uppdelat: "läsa" står sist i sammansättningen och är huvudverbet

M7102-620070. Utgåva 1: 1970 april. 1000 bl à 100 ex (FOA 604)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard | EBCDIC | BCD | Speciell stans-instruktion | Tecken | Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel
DATA / ALGOL / PL/I: Sats / Sats | Identifikation / Sekvensnr

1–8	9–10	11–14	15–20	27–	66–73
0407000313		1-14	2	Litteraturförteckningar	50
	14	1		som (litteraturförteckningar)	53330
	15			dyker upp	40
	16			med jämna mellanrum	43
	17			i postfacken	44
	18			också	42
40	1	1-11	1	Eftersom	462
	2			jag	30
	3			har	40
	4			goda	45
	5			relationer med	40
	6			bibliotekspersonalen	501
	7	1 4	2	kan	40
	8			jag	30
	9			verkligen	42
	10			inte	--
	11			klaga på	40
	12			servicen	50
	13			(från biblioteket)	53
50	1			Det	99
	2			är	99
	3			den (biblioteksservicen)	30
	4			som	99
	5			är	41
	6			viktig	32

Annotations:

modus bryts

orsaks-sats

verbfras - adj. blir gradadverb

andra satsen är negerad

emfatisk som-sats

M7102-620070. Utgåva 1: 1970 april. 1000 bl à 170 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare			Teckenstandard	EBCDIC / BCD	Speciell stansinstruktion	Sida av
Kundnummer	Telefon nr	Datum				Program

Assembler language: Namn 8 · 10 Operation 14 · 16 Operand · Kommentar · Identifikation / Sekvens
FORTRAN: Fml nr 5 · 6 · 7 FORTRANformel
DATA
ALGOL Sats
PL/I Sats

1–10	11–14	17	24	26–	66–	71–	73–
0508000101	1124	1	1	Jag	30		
2				ville	40		
3				inte	--		
4				sticka(ge sig i våg)	40		
5				till den där	61		
6				trista	62		
7				hålan(Kalmar)	60		
8		2		trots att	462		
9				många		31	
10				(personer)		30	
11				hade sagt		40	
12				(till mig)		70	
13	2 1			att		503	
14				det(Kalmar)			30
15				faktiskt			42
16				var			41
17				riktigt trevligt			32
201	1-21			Jag	30		
2				frågade mig	40		
3			1	vad	502		
4				lärarkandidaterna		30	
5				hade tänkt sig		40	
301	1-11			Dom(lärarkandidaterna)	30		
2				undrar	40		
3				ofta	43		

"vilja" = hjälpverb = en sats

medgivande

supplering av tvetydigt verb - "neutral" form

Objekt(Mål)

ny källa till påståendet

supplering av obest. huvudord på ny rad

obs tempus "var" = "är"

indirekt fråga

M7102-620070. Utgåva 1 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard: EBCDIC / BCD | Speciall-stans-instruktion | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar

FORTRAN: Fml nr 5 | 6 | 7

DATA, ALGOL Sats, PL/I Sats

Identifikation / Sekvens

källa

direkt fråga

Kolumn 1–10	Kolumn 11–14 (källa)	Kolumn 21 (direkt fråga)	Kolumn 25–	Kolumn 66–71
0508000304	2-11	1	vad	502
5			har	40
6			vi (lärarkandidater)	30
7			för nytta av	40
8			det här (pedagogik)	50
0601000101	1-21		jag	70
2			utsattes för	40
3			hård	52
4			kritik	50
5			(av lärarna)	30
6	3		utan att	462--
7			man (lärarna)	30
8			kunde precisera	40
9			tillräckligt	45
10			för mig	70
11			vad (fel)	50
12			som	99
13			var	99
14			fel	88

subjunktions-sats = negerad

en lösning på den sista satsen kan vara att låta huvudordet bli supplerat in efter "vad". Kod 99 tar hand om två ord och vi tillgriper den tekniska koden 88, som säger att ordet ej räknas på detta ställe utan finns i annan kod i meningen

M7102-620070. Utgåva 1. 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard: EBCDIC / BCD | Speciell stans-instruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar | Identifikation / Sekvens
FORTRAN: Fml nr 5 | 6 | 7 FC
DATA
ALGOL Sats
PL/I Sats

Kod (kol. 1–16)	Operand (kol. 25–)	Kol. 66–71
07010001011-241	Om	462
2	uppdraget	30
3 21	(att inventera problem inom)	33
3 22	(yrkespedagogik)	33
4	hade gått i lås	401
5 2	så	99
6	hade	40
7	jag	30
8	haft	40
9	ett projekt	50
10	att satsa på	53
2011121	Jag	30
2	visste	40
3	inte	--
4	vad för infallsvinkel	50
5 -	(som)(för infallsvinkel)	53250
6	jag	30
7	skulle ha	40
8	för infallsvinkel	88
301 1-21	Sö	30
2	visste	40
3	vilken(infallsvinkel)	50
4 3	som(infallsvinkel)	53230
5	vore	41
6	lämpligast	32

Annotations:

villkor

enhet över två rader

ord-nummer

idiom = ej supplering

verb som bestämning

frågekod behövs ej

negering slut

ett annat fall av 88-kodning - verbet = "känna till" + Obj + relativsats

vid 88 - ingen supplering

konj. form på kopulan

vanlig supplering

A7102-A20070. Utgåva 1. 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard | EBCDIC / BCD | Speciell stans-instruktion | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel
DATA, ALGOL Sats, PL/I Sats | Identifikation / Sekvensnr

1 5 10 15 20 25 30 35 40 45 50 55 60 65 71 73 80

Några ytterligare flätningsexempel:

Kol. 11	Kol. 26–	Kod
2	Pedagogiken	88
	ansåg	40
	dom (lärarkandidaterna)	30
	(att)	502
	pedagogiken	30
	vara	41
	ett teoretiskt	32
	komplex	30
/	Tråkig	32
	tyckte jag	42
	att	99
	han (MN)	30
	var	41
	Jag	30
	läste	40
	nån bok	50
	som (bok)	53288
	jag	30
	visste	40
	(att)	503
	boken	30
	var	41
	bra	32

källkod

jämför!

att-insättning

"att" blir innehållslöst

flätnings-koden i bisats

M7102-620070. Utgåva 1: 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Specialstansinstruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 — Kommentar

FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel — DATA — ALGOL Sats — PL/I Sats — Identifikation / Sekvensnr

Några exemplifieringar från kapitel 5:

Kol. 1–25	Kol. 26–	Kol. 66–68	Kol. 69–71
1-21	Jag	30	
	såg	40	
	(att)	502	
	honom(NN)		30
	sitta		40
	i biblioteket		44
1-21	Jag	30	
	bad	40	
	honom(NN)	70	
4	(att)	502	
	(han)(NN)		30
	(skulle skriva) skriva		40
	en idé'skiss		50
1-11	(Att)	302	
	X:s		30
	ideliga		45
	sänkningar		40
	av forskningsstandarden		501
	droar	40	
	stora	52	
	grupper	50	
	(av forskare)	53	

M7102-620070. Utgåva 1. 1970 april. 1000 bl à 100 ex
(FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard | EBCDIC | BCD | Speciell stans-instruktion | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 | Kommentar

FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel

DATA ALGOL Sats PL/I Sats

Identifikation / Sekvensnr

Den kodade texten
från ruta 1:

1–10	12–15	25–26	27–	66–67
1106000101	1-11		Jag	30
2			är	41
3			en väldigt dålig	32
4			forskare	30
5		21	när det gäller just sånt där	46
5		22	(informationssökning)	46
201	1-11		Men	99
2			jag	30
3			är	41
4			annars	99
5			väldigt osystematisk	32
6		21	i hela mitt sätt att söka	46
6		22	(informationssökning)	46
301	1121		Och	99
2			tex	99
3			jag	30
4			har	40
5			aldrig	--
6			gått på (sökt)	40
7			någon sådan här	51
8			referatsammanställningar	50
9			nån gång	43
401	1-21		Men	99
2			däremot	99
3			så	99

M7102-620070 Utgåva 1: 1970 april. 1000 bl à 100 ex
(FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard: EBCDIC / BCD | Speciell stans-instruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel
DATA ALGOL Sats, PL/I Sats | Identifikation/Sekvens nr

1–10	11–14	20	Operand (25–)	66–72
1106000404	1-21		iblan d	43
5			har	40
6			jag	30
7			ägnat mig åt	40
8			systematik	50
9			när det gäller att	532
10			(jag)	30
11			leta igenom	40
12			tidskrifter	50
13			och	++
14			sådant där (information)	50
501	1-11	1	Men	99
2			den (systematiken)	50
3			överger	40
4			jag	30
5			snabbt	45
601	1-11	2	Medan	99
2			jag	30
3			däremot	46
4			punktmässigt	45
5			läser in mig	40
6			på ett område (forskningsområde)	50
7			eller	++
8			en tidskrift	50
9			väldigt väl	45

M7102-620070. Utgåva 1 1970 april. 1 000 bl à 100 ex (FOA 304)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Tecken-standard: EBCDIC / BCD | Speciell stans-instruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel
DATA
ALGOL Sats
PL/I Sats 5
Identifikation / Sekvensnr

1–10	11–15	20	27–	67–68
1106000701	1-21		Jag	30
2			har	40
3			dels	99
4			utnyttjat	40
5			väldigt lite	45
6			den typen av	51
7			sammanställningar	50
8			abstracts	50
9			och sådant	99
801	1-11		Och dels så	99
2			har	40
3			jag	30
4			väldigt få	51
5			rekommendationer	50
6			(för informationssökning)	53
901	1-21	1	Vete fan om	42
2			det	99
3			inte	42
4			är	99
5			nån (forskare)	30
6			som	99
7			har	40
8			nån gång	43
9			skrivit om	40
10			det där (referatsammanställningar)	50

M7102-620070 Utgåva 1. 1970 april. 1000 bl à 100 ex (FOA 404)

PROGRAMMERINGS- OCH DATABLANKETT

Programmerare | Kundnummer | Telefon nr | Datum | Teckenstandard: EBCDIC / BCD | Speciell stansinstruktion: Tecken / Stansning | Sida av | Program

Assembler language: Namn 8 | 10 Operation 14 | 16 Operand | Kommentar
FORTRAN: Fml nr 5 | 6 | 7 FORTRANformel
DATA
ALGOL Sats
PL/I Sats
Identifikation / Sekvensnr

Kol. 1–10	Kol. 11–14	Kol. 15–20	Kol. 25–	Kod
1706000911	1-11	1	att	502
12			det	99
13			är symtom på	40
14			en intellektuell	52
15			kris	50
16			alltså	42
17		4 1	att	303
18			man	30
19			måste banta ner	40
20			information	50
21		2	för att	464
22			(man)	30
23			kunna ta emot	40
24			den (informationen)	50
25			alltså	42
1001	1-14	2 1	För att	462
2			man	30
3			ska kunna använda	40
4			information	501
5		1	så	99
6			måste	40
7			man	30
8			förstöra	40
9			den (informationen)	50
10			alltså	42

L47102-620070. Utgåva 1: 1970 april. 1000 24 à 100 ex (POA 404)

8. VAD GER DATORN SVAR PÅ ?

Om vi förutsätter att kodningsreglerna har tillämpats på samma sätt hela texten igenom, kan vi säga hur våra koder ser ut, dvs vilken information vi kan hitta på ett visst ställe i texten. Vi kan "beställa" listor med innehållet i våra paradigmatiska koder och studera varje kort (utskrift), dels för att kontrollera felstansningar och felkodningar, som sedan kan rättas med hjälp av identifikationskoderna, dels för att göra strukturbeskrivningar av de olika elementen, som kan bilda grunden för datorbaserade analyser. Vid beskrivningar av strukturer analyseras varje segment, som får en beteckning, t ex ifråga om syntaktisk position, morfologisk bestämning eller ordklasstillhörighet, beroende på hur djupgående analysen ska bli. Om vi vill att datorn ska utföra vissa instruktioner avseende vissa segment, t ex kontrollera vilka prepositioner som föregår vissa substantiv inom vissa koder, måste den först veta vad substantiv är, respektive vad prepositioner är. Den informationen specificeras i särskilda lexikon, som datorn "matchar" texten mot. Ett stort problem i det här sammanhanget är frågan om homografi, dvs när ett ord trots flera betydelser stavas lika. Problemet uppstår särskilt när man bildar lexikon, där ett ord endast ska förekomma en gång i grundform. Det betyder t ex att olika ändelser "skalas bort". När texten söks igenom med avsikten att hitta, registrera eller manipulera ord, jämförs ordet med den stam som återstår efter att de specificerade ändelserna i ett ändelselexikon har uteslutits. Vi ska inte här gå igenom principerna för hur detta kan åstadkommas, framförallt därför att homografiproblemet inte ännu är tillfredsställande löst inom den datalingvistiska forskningen. Här ska endast nämnas att vi konstruerat fyra slag av lexikon, (1) ändelselexikon, (2) substantivlexikon, (3) verblexikon och (4) adjektivlexikon. Redan denna struktur tillåter frågor till datorn, som vi kan få en hel del information av, t ex frekvenslistor avseende ord eller koder. Sedan är det möjligt att genom kopplingar, t ex med hjälp av Boolsk algebra, strukturera materialet så som vi just för tillfället önskar. Den möjligheten har vi med våra många differentierande övergripande koder i kombination med de övriga. För att läsaren eller användaren av manualen ska inse dessa möjligheter, ska vi i det följande ställa ett antal frågor, som vi besvarar genom att ange på vilket sätt kopplingar och differentieringar skulle kunna ske, dvs vilka instruktioner datorn får och vilken "kunskap" den förutsätts använda. (Vi använder inte formelspråk här.) Inferenser kan vi inte åstadkomma i det här sammanhanget, eftersom sådana förutsätter statistiska beräkningar.

Fråga: Hur värderas det som personerna talar om och hur många värderande uttalanden görs?

Svar:

Vi vet att personerna uttrycker sig med beskrivande adjektiv till objekt. Vi har skalvärden på adjektiv och verb registrerade i lexikon, som båda uttrycker en värdering av de objekt man talar om. Men värderingar kan vara förtäckta (implicita). Vi kan gå tillväga på flera sätt här.

(1) Vi ger datorn instruktionen att leta upp alla verb (kod 40) och alla adjektivattribut (koderna 32, 52, osv). Vid varje utfall i sökningen på dessa koder ska ordet kontrolleras mot verb- respektive adjektivlexikonen, där skattningar är angivna för varje element. Datorn tar hjälp av ändelselexikonet för att hitta grundordet. Finns det, så återförs dess skattade värde till texten och kopplas med substantivet (t ex i objektkoden). En summering görs av objekten, varvid substantivlexikonet används, och värdet av substantivet införs vid dess plats i lexikonet. Sedan kan vi få sorterat hur värderingen ser ut beroende på om substantivet är använt som objekt och vilket slag, samt om det har använts som agent.

(2) Vi kan även anta att "värderingar" enligt frågan gäller just värderande omdömen (explicita), uttryckta genom 'är'-liknande konstruktioner. Instruktionen gäller då: Skriv ut alla predikativ, där verbet har kod 41 och där det predikativa ledet som följer efter 41 innehåller endast 32 eller 32 i kombination med 30.

(3) När det sen gäller antalet värderingar avser instruktionen en beräkning av proportionen 41-satser i förhållande till övriga satskonstruktioner, dvs alla satser med 40. Sats definieras då som varje slinga (i satskolumnerna) som innehåller 40. Om sats skulle definieras som "mening" används identifikationskoderna, dvs varje mening börjar, där i kolumn 9-10 finns koderna 01.

Fråga: Hur många säkra påståenden gör personerna?

Svar:

Datorn instrueras att söka på kolumnerna 14-23. Har kolumn 14 kodats med 1 gäller meningen eller satsen som "säker", förutsatt att övriga temakolumner är tomma. När detta är uppfyllt kan man genomsöka satskodningen på kod 42. Det förutsätts då att ett lexikon är bildat, som anger om uttrycket tillhör en kategori "säker" (i förhållande till "osäker", "trolig", osv), när antagande inte kodats i kolumn 22. De satser som genom dessa specificeringar av "säkerhet" skrivs ut kan sedan sättas i relation till antalet övriga satser.

Fråga: Hur mycket är refereringar till händelser bakåt i tiden respektive framåt i tiden?

Svar:

Här är den primära informationen förlagd till kolumn 13. Finns där kod 2, anger den att något skett, eller förväntades ske vid en tidigare tidpunkt än den närvarande, medan kod 3 anger en kommande tidpunkt. De operationer man önskar få utförda gäller således endast med dessa förutsättningar. För att vara säker på att kodningen i kolumn 13 inte också avser sådant som inte faktiskt hänt, specificerar man att kolumn 14 skall innehålla kod 1.

Fråga: Vilka villkor ställer intervjupersonerna för att deras utsagor är giltiga?

Svar:

Giltiga utsagor gäller sådana som är utan förbehåll. (Vi utgår från att det inte gäller den filosofiska frågan om 'sann' och 'falsk'.) I vilka koder kan det finnas förbehåll? Instruktionen bör i första hand vara att beakta satser med kod i kolumn 15 (villkor). Förbehåll kan också uttryckas modalt, och vi måste därför ange att kod 4 i kolumn 14 gäller här. Förbehåll kan dessutom finnas i funktionskod 46. Villkor kan vi därför spåra genom att sammankoppla kolumn 15 (villkorskod 1) + satser med kod 4 i kolumn 14 + satser med kod 46, som samtidigt inte har temakod (i kol 15-23).

Fråga: Hur många kontradiktoriska uppgifter finns det egentligen?

Svar:

Här kunde vi ha möjlighet att utnyttja den meningsövergripande kodningen, dvs instruktionen blir att räkna eller skriva ut de meningar som i kolumn 19 har kod 1 eller 2, varvid de skrivs ut i ordningsföljd. Vi kan dessutom utnyttja vikterna i adjektiv- och verblexikonen, på så sätt att de ord eller objekt som värderas kopplas med respektive persons sammansättning av dem i värderande utsagor. Om t ex ordet "forskare" värderas dels negativt, dels positivt, är det en indikation på att personen motsäger sig själv. Ett autentiskt exempel (skalan anger 1 = neg, 7 = pos):

forskare & oerfaren = värderas 2.75
forskare & öppen = värderas 5.50

Genom kolumn 19 utvinner vi hela meningar och genom att utnyttja de viktade lexikonen utvinner vi specificerade omdömen om enskilda begrepp. Att det gäller en och samma person anges genom identifikationskoderna i början (kol 1-2).

Fråga: Hur mycket förutsätter intervjupersonerna att intervjuaren förstår av vad som sägs, dvs hur mycket "ekonomiserar" man sina meddelanden?

Svar:

Det här är ett problem som gäller våra suppleringar. Vi skulle kunna beräkna några index här. En instruktion kan vara att beräkna proportionen supplerade enheter (dvs enheter inom parentes) i förhållande till den text (de enheter) som i realiteten har sagts. En annan instruktion kan gälla beräkningen av hela in-supplerade satser i relation till det totala antalet satser.

Vad man här bör försöka göra är att bilda ett mått som säger något om naturliga referenser (t ex olika slag av pronomen) i förhållande till "kända" och "okända" begrepp, dvs vi får ställa frågan: Vad är naturligt att utelämna i förhållande till vad som redan sagts, och hur har man valt att utnyttja detta egentligen? Här måste vi alltså studera suppleringarnas art, syntaktiska och morfologiska funktion, avstånd till varandra (vid samma supplering) eller liknande. En hypotes kan vara att det man objektiverar eller kommenterar (objekt) i proportionellt högre grad är definierade än det man talar om (subjekt). Det skulle då bli en fråga om en s k topic-comment-analys, som är syntaktisk och där vi då utnyttjar de olika ledens upprätthållna ordningsföljd i meningen. Allt detta är "konstruktionsarbete", vars regler inte kan anges på det här stadiet.

Dessa frågor och svar bör vara tillräckliga för att visa hur koderna principiellt fungerar. Identifikations-, tema- och funktionella koder (enstaka eller i satskolumner) är samtliga nödvändiga och kan komplettera varandra bra. Än har vi inte kommit så långt att vi kan ange var brister finns. Men vi kan säga att de kan fungera på det sätt som skisserats, förutsatt att vi specificerar olika slag av lexikon, vilket för övrigt är nödvändigt vid alla slag av textanalyser. Lexikonen konstrueras i förhållande till uppkomna syften och kan således fungera på ett mycket flexibelt sätt. Flera av de ovan nämnda frågorna har i realiteten ställts och vi har kunnat göra intressanta analyser med hjälp av statistiska metoder.

Vår föreslagna innehållsanalysteknik utgår ifrån en modell. För att ge bra svar ur materialet, krävs att frågorna är bra, dvs frågorna måste vara relevanta utifrån någon teori eller modell för vad som i det enskilda fallet angetts som "innehåll". Om man har en sådan innehållsanalysmodell och har matat in relevanta enheter, bör inga irrelevanta svar återvinnas, men de ska tolkas.

9. INSKOLNING AV KODARE, TEST OCH ÖVERENSSTÄMMELSEKONTROLLER

Det program för inskolning och träning av kodare som ges här följer de regler som har angivits i de föregående kapitlen. De förslag till hur programmet på lämpligaste sätt kan genomföras har tillkommit genom erfarenheter från utprövningar av reglerna.

9.1 Förutsättningar

Innan en träningsperiod genomförs ska den som ansvarar för undersökningen ha gjort klart vad som gäller ifråga om metod, material och modell, dvs de frågeställningar som behandlades i kapitel 1.

9.1.1 Provmaterial

Den text på vilken träningen ska ske ska vara utskriven, lämpligen på A4-ark. Det kan också rekommenderas att denna provtext är av samma slag som den text som ska analyseras. Vid en intervjuundersökning kan ett par provintervjuer genomföras, vars texter då utgör träningsunderlaget.

9.1.2 Inlärningssituation och hjälpmedel

En inskolning i användandet av ett omfattande regelsystem som detta kräver en omsorgsfull planering. Arbetet sker stegvis med stigande svårighetsgrad och uppföljning av de olika inlärningspassen bör ske i så nära anslutning till lektionerna som möjligt. Uppföljningen ska innefatta repetition av inlärda regler och rättning av provtexter. Denna situation kräver personliga instruktioner och individuell hjälp av en "lärare", som är väl förtrogen med reglerna och materialet.

Samtliga kodare är närvarande under lektionerna. Antalet lektioner är beroende av kodarnas snabbhet att tillägna sig reglerna. Ett krav är dock att lektionerna följer tätt på varandra. Kontinuitet är viktigt för att de olika stegen ska upplevas som en integrerad helhet efterhand. Att arbeta med längre textavsnitt är dessutom tidsödande och kräver avskildhet och avspänning.

Lektionerna bör hållas på den institution där undersökningen bedrivs dels därför att tillgång till tekniska hjälpmedel finns vid behov, dels därför att hjälpmedel som t ex ordböcker är lättillgängliga. Dessutom finns ansvarig personal (projektpersonal) tillhands.

De hjälpmedel som kan komma till användning är, förutom kodningsmanualen, t ex SAOL för kontroll av korrekt stavning, likaså någon ordbok för främmande ord eller facklexikon. Man kan nämligen inte utesluta att vissa begrepp har missförståtts av den personal som skrivit ut texterna. Av samma anled-

ning är det också viktigt att ha tillgång till amerikanska eller engelska ordböcker, eller sådana övriga hjälpmedel som är relevanta mot bakgrund av den undersökning man arbetar med (jfr våra standardiserade suppleringar).

9.2 Urval av kodare

Det finns många hänsyn att ta vid urval av personer som ska tränas att utföra ett omfattande kodningsarbete med största möjliga tillförlitlighet. Urvalet måste grunda sig på dels personliga förutsättningar, dels praktiska ifråga om tidsåtgång och ibland även årstid. Det senare beror dock på vilken planering undersökningsledaren avser att hålla. Ur alla synpunkter är arbete på dagtid att föredra, för alla parter.

Angående personliga förutsättningar ska följande anföras:

1. En första förutsättning är intresse och vilja till samarbete. Förståelse och inblick i detta slag av kreativt arbete är en stor fördel.
2. En god allmänutbildning är ett krav. Det har visat sig att utbildningsnivån bör vara akademisk grundexamen, språklig eller samhällsvetenskaplig. En studerande med starka intressen i båda riktningarna kan också vara lämplig.

När man bestämmer utbildningsnivån ska man ta hänsyn till det stoff som textmaterialet innehåller (i vårt fall pedagogik-psykologi). Det kan inte tas för givet att det behandlade ämnet kan förstås av alla, oavsett utbildning och kunskaper. Men det kan också vara så att en i samhällsvetenskapliga eller beteendevetenskapliga frågor mindre insatt person kan klara av kodningen enbart med hjälp av sin språkliga kunnighet, som de flesta regler bygger på. Våra utprövningar har visat att kunskapsbrister i språkstruktur inte kan uppvägas av kunskaper i eller kännedom om analystextens ämne. Insatt i ämnet kan man bli efterhand som arbetet fortskrider.

3. Kodare bör vara uthålliga, envisa och vara stabila till sin läggning. När man tror sig ha funnit de mest lämpade personerna (minst två) ska först vissa baskunskaper prövas, som är relevanta för denna analys

Ett första urval kodare för förkunskapstest kan innefatta en större grupp, varifrån man har möjlighet att välja ut dem som har de bästa förutsättningarna att fortsätta.

9.2.1 Kontroll av förkunskaper och tillförlitlighet

Här framläggs inget färdigt prov, som kan ges generellt. Varje instruktör bör kunna avgöra vilken text som i det enskilda fallet är lämpligast eller på vilket annat sätt testet ska ske. Förtestets karaktär gör emellertid att skriftspråkstext bör användas och helst med anknytning till analysens ämnesområde. Ett förslag som lämpar sig för vårt syfte ges i bilaga 3.

Följande förkunskaper ska testas. Kodarna ska

1. kunna identifiera ordklasserna substantiv, verb, adjektiv, pronomen och adverb
2. kunna identifiera meningar och satser
3. kunna identifiera satsers olika syntaktiska enheter, dvs sådana enheter som i traditionell satsanalys kallas subjekt, predikat, predikatsfyllnad, objekt, attribut och adverbial

Om dessa kunskaper inte tillfredsställande har redovisats av presumtiva kodare, rekommenderas instruktör eller undersökningsledare att söka andra lämpliga personer. Problemet blir annars för stort med inskolningen och det tar lång tid, vilket ofta är ett dilemma vid tidsspecificerade undersökningar. En person som inte har denna bas har svårt att inom rimlig tid lära sig och förstå reglerna i manualen. Riktlinjer för hur man bör bedöma detta test ges i bilaga 3:3-3:5.

Om kunskaperna är tillfredsställande vidtar ett test på provmaterialet, dvs den autentiska analystexten. Detta prov ska testa om kodarna kan

1. tillämpa de ovan beskrivna baskunskaperna på autentiskt material

Vi har erfarenheter av att man kan klara av förkunskapsprovet utan att för den skull kunna tillämpa samma kunskaper på nedskriven taltext. I det här sammanhanget lär inte ordklasserna vålla bekymmer. Det är svårare när det gäller satser och satsdelar, eftersom texten kan innehålla många ofullständiga satser och inte behöver vara interpunkterad som "normal" skriftspråkstext.

Även här ges ett förslag till provtext, hämtad ur vårt material. Testet ska ges utan föregående inskolning, så därför har uppgifterna anpassats efter textens utseende. Denna text tillsammans med rättningsanvisning ges i bilaga 4. (Några introducerande ord om textens ursprung och innehåll bör dock sägas av instruktören. Man kan också tala om att den fortsatta träningen gäller just denna typ av text.) När även detta test rättats, utgör de två testen underlag för urval av (minst) två personer som ska fortsätta det större kodningsarbetet.

Här är det viktigt att veta om två personer har svarat likadant på testet, dvs inte bara åstadkommit lika antal poäng utan också lika svar på en och samma uppgift. Mätningen avser i det fallet överensstämmelsen mellan två av varandra oberoende kodare (se kap 9.4). Om överensstämmelsen inte är tillförlitlig kan

man även här välja nya kodare som då genomgår alla test från början. Man kan också välja att träna denna senare fas med samma kodare och testa på nytt. Liknande texter gås då igenom, där det anges hur tillämpningen av kunskaper ska ske. Det senare kan förordas, eftersom det kan sägas tillhöra den fas då man bör presentera de första delarna i manualen. Denna fas innebär då en samträning. Här sker också en övergång från baskunskaperna till kodningsreglerna. Den stegvisa träning som beskrivs under 9.3.2 tar vid. Varje steg kan avslutas med test. Steg 2 och 3 kan också lämpligen delas in i mindre avsnitt med test mellan varje. När kodarna tränats på samtliga steg görs ett test, där kodarna ska

2. korrekt mot mall tillämpa kodningsregler

Denna kontroll görs för att testa om reglerna förstås på samma sätt av kodarna. Mätning sker genom en summering av antal korrekt kodade analysenheter ställda i relation till det totala antalet analysenheter i testet. Dessutom granskas överensstämmelsen mellan två av varandra oberoende kodare (enligt kap 9.4).

Om resultatet av dessa test inte uppfyller de uppställda kraven, tas ställning till om kodarna ska bytas ut (således börja från början med initialtestet) eller om kodningsreglerna bör förtydligas på några punkter. Om detta är fallet, skrivs den eventuella ändringen in och nytt test görs ifråga om tillämpningen mot mall följt av ny mätning tills tillfredsställande resultat har nåtts. Därefter följer den egentliga kodningen. Om den tar lång tid i anspråk är det lämpligt att med jämna mellanrum testa på nytt, eftersom kodarna kan få ny referens till reglerna efterhand, och nytillkomna problem i texterna kan göra att en regel borde förtydligas.

9.3 Instuderingsschema

9.3.1 Målbeskrivning

Det program som ges nedan har utformats på grundval av utprövningar och praktiska erfarenheter av materialet och dess karaktär. Målsättningen är att två eller flera kodare ska kunna urskilja analysens enheter och tillämpa kodningsinstruktionerna på ett likartat sätt. Därför är det viktigt att instruktören går tillväga på ett så adekvat sätt som möjligt med hänsyn till analysens målsättning. Mot bakgrund av de förutsättningar som angetts i kapitel 1 om metod, material och modell, bearbetningen av text enligt anvisningar i kapitel 2 samt tillordningen av koder i kapitel 4 har följande delmål uppsatts för kodarna:

1. utifrån den i intervjun ställda frågan avgränsa relevant textmaterial från icke relevant

2. segmentera texten i meningsfyllda enheter enligt i förväg fastställd definition av "mening"
3. identifiera de syntaktiska relationer som råder mellan enheter inom varje segmenterad mening och sats och därvid supplera icke explicita enheter, så att analysens paradigm görs fullständigt enligt i förväg fastställda kriterier
4. enligt anvisningar skriva ner på datablad de enheter inom varje segmenterad och supplerad mening eller sats som skall ingå i analysen
5. särskilja olika satstyper enligt definition i kodningsmanual
6. klassificera de nedskrivna analysenheterna enligt deras syntaktisk-funktionella innehåll
7. tillordna koder i därför avsedda kolumner på datablankett

9.3.2 Stegvis träning

En träning till fullständig behärskning av kodningsreglerna ska ske på så sätt att de olika stegen följer på varandra i en logiskt meningsfull sekvens. Vi har funnit att det är lämpligt att dela in träningen i tre steg:

Steg 1. avgränsa relevant text och segmentera texten i meningar

Steg 2. identifiera syntaktiska relationer, supplera till fullständigt paradigm och skriva på datablad

Steg 3. särskilja satstyper, klassificera analysenheter och tillordna koder.

Avsnitten i manualen är för de olika stegen (1) kapitel 1 som introduktion, men främst kapitel 2 och ruta 1 som utgångspunkt, (2) kapitel 3, 5 och 6 samt exempelsamlingen (i kapitel 7) och (3) kapitel 4 och 7.

När en viss vana har uppnåtts går de olika stegen in i varandra och uppskrivning på bladet sker direkt. Innan man nått så långt torde dock det andra steget vara mest tidskrävande, eftersom momentet 'skriva' innefattar många tekniska aspekter, som måste tränas till en automatik.

9.4 Beräkning av överensstämmelse

Flera utprövningssteg har föregått den här manualen och vi har bl a utarbetat riktlinjer för hur överensstämmelsen kan protokolleras och vilka beräkningar vi gjorde. Vilka analysenheter som medtages eller slås samman i kategorier beror på hur kodningsreglerna är strukturerade, dvs vad som är av samma typ. Reglerna tillåter i vårt fall följande kategorier:

1. Avgränsning av meningar och satser.
2. Kodning av källa, tempus, modus (som alltid förekommer).
3. Avgränsning, supplering och kodning av textenheter (begrepp) som i sin tur kan delas i huvudbegrepp (enligt paradigmet) och övriga.

När det gäller att föra protokoll över kodningen måste rigoröst utformade an-

visningar göras för att en statistisk behandling ska kunna ske. Vi gjorde olika slag av markeringar på databladen, vilka sedan summerades och var underlag för beräkningar. I bilaga 5 ges de anvisningar vi utarbetade.

Vid den statistiska bearbetningen, där vi utgått från det dikotoma förhållandet överensstämmelse - icke överensstämmelse, använde vi först Osgoods index för kodaröverensstämmelse

$$\frac{2\ (\text{gem } A + B)}{(\text{tot } A) + (\text{tot } B)}$$, där

(gem A + B) anger antalet gemensamma bedömningar för kodare A och B

(tot A) anger det totala antalet bedömningar kodare A har gjort, dvs summan av de bedömningar som är gemensamma med kodare B och de bedömningar som inte överensstämmer med kodare B:s bedömningar

(tot B) anger det totala antalet bedömningar som kodare B har gjort

Vid perfekt överensstämmelse av t ex 150 bedömningar är överensstämmelsen

$$\frac{2 * 150}{150 + 150} = 1.00$$

Osgoods index är inget direkt mått på överensstämmelsens storlek. Ur detta mått kan man avläsa skillnader i svårighet att bearbeta olika frågor och intervjuer och det visar förändringar i överensstämmelsen vid olika tidpunkter. Man bör observera att Osgoods index inte är detsamma som den procentuella överensstämmelsen; det ger något lägre värden.

Vi satte 80% överensstämmelse som minimikrav. (Det bör åtminstone vara omkring .70.) För att undersöka om de funna värdena uppfyller detta kriterium används lämpligen binomialtestet, som är särskilt lämpligt vid dikotomt fördelade data, dvs här överensstämmelse respektive avvikelse mellan kodare. H_0: överensstämmelse mellan kodarna är .80 eller högre, testas mot H_1: överensstämmelse mellan kodarna är lägre än .80. Ett z-värde beräknades med följande formel:

$$z = \frac{(x \pm 0.5) - NP}{N * P * Q} \quad \text{om } N * P * Q \geq 9, \text{ där}$$

x anger antalet gemensamma bedömningar för kodare A och B

$\pm$ 0.5 anger en konstant för kontinuitetskorrektion

N anger det totala antalet bedömningar

P anger kriteriet, .80

Q = .20

z-värdet testas på 5%-nivån. Värden under (till vänster om) -1.64 tyder på att 80% överensstämmelse inte har kunnat säkerställas.

I de fall $N * P * Q \geq 9$ har den exakta sannolikheten använts (Tables of the cumulative binomial probability distribution, 1955). P-värden under .05 tyder på att kriteriet .80 inte har uppnåtts.

Vid beräkningar inom de olika kategorierna på ett större testmaterial uppnådde vi höga värden på överensstämmelse, de flesta omkring .90 och däröver. Dessa värden baserade sig då på (1) avgränsning och (2) kodning av satser, (3) avgränsning och (4) kodning av enheter (begrepp). När vi delade upp de olika enheterna i huvudord (inkl. verb), attribut och adverb, visade det sig att vissa adverb hade lägre överensstämmelse, samt det efterställda attributet. Av detta skäl har den här versionen av manualen förtydligats på dessa punkter.

Höga värden uppnås ännu lättare om man gör beräkningen enbart på den gemensamma texten. Man utesluter då avvikande begrepp, (t ex orsakade av suppleringar och strykningar). Denna gemensamma text är då lik skriven text, som är det hittills vanligaste utgångsmaterialet vid innehållsanalyser.

9.4.1 Bestämning av vad som är avvikelse

Avvikelserna hos två kodare kan bedömas olika beroende på vikten av dem. Detta avgörs i förhållande till analysens syfte och de manipuleringar man avser att få utförda maskinellt. Det är t ex i vårt fall viktigt att i första hand AaO-paradigmet är rätt kodat och de attribut och bestämningar som hör till. Vissa partiklar måste höra till verbet, andra till substantivkoderna, osv. I och med att vi använder både adjektiv och verb som bestämmare (viktare) till substantiv kan man däremot tänka sig att predikativ (adjektiv -41+32 eller verb -40) kan betraktas likvärdigt. Huruvida man "drar ut" en relativkonstruktion i ny satskolumn (när den är restriktiv, av typen "den information som finns tillgänglig") eller sätter den som en enhet efter huvudordet är också en diskutabel avvikelse. Det beror på vilken betydelse man tillmäter verbet och om man önskar skilja på den strukturen och andra.

En stor betydelse har enheternas utseende på varje hålkort. När man vid konstruktion av olika slag av lexikon är angelägen om att en viss kod är så "ren" som möjligt, tillmäts detta stor vikt, även om det gäller koder utanför de huvudsakliga.

Det är alltså viktigt att man avgör sådant före protokolleringen mot bakgrund av vilken roll de olika enheterna spelar i innehållsanalysen.

10. REFERENSER

Abelson, R.P. The structure of belief systems. I: Schank, R.C. & Colby, K.M. Computer models of thought and language. San Francisco: Freeman, 1973, Ss 287-339.

Allén, S. Indelningen av nusvenskans pronomen. Arkiv för nordisk filologi, 1958, 73, 51-70.

Allén, S. On phraseology in lexicology. (Göteborg: Avdelningen för språklig databehandling), 1976.

Allén, S. Text-based lexicography and algorithmic text analysis. (Göteborg: Avdelningen för språklig databehandling), 1976.

Bjerstedt, Å. Den pedagogiska processen. (Pedagogisk orientering och debatt, 31.) Lund: Gleerup, 1970.

Brodda, B. (K)overta kasus i svenskan. PILUS(Stockholm: Institutionen för lingvistik), Nr 18, 1973.

Diderichsen, P. Elementaer dansk Grammatik. (3:e uppl.) Köpenhamn: Gyldendal, 1962.

Engelkamp, J. Semantische Struktur und die Verarbeitung von Sätzen. Stuttgart & Wien: Hans Huber Verlag, 1973.

Fillmore, C.J. The case for case. I: Bach, E. & Harms, R.T. Universals in linguistic theory. New York: Holt, Rinehart & Winston, 1968.

Fillmore, C.J. Lexical entries for verbs. Foundations of Language, 1968, 4, 373-393.

Grieve, R. Definiteness in discourse. Language and Speech, 1973, 16(4), 365-372.

Halliday, M.A.K. Language structure and language function. I: Lyons, J. (Ed.) New horizons in linguistics. London: Penguin Books, 1970.

Halliday, M.A.K. Functional diversity in language as seen from a consideration of modality and mood in English. Foundations of Language, 1970, 6, 322-361.

Harris, R.J. Effects of nonlinguistic knowledge on language production. Journal of Psycholinguistic Research, 1974, 3(4), 303-310.

Holsti, O.R. A computer content-analysis program for analyzing attitudes: The measurement of qualities and performance. I: Holsti, O.R. Content analysis for the social sciences and humanities. Reading, Mass.: Addison-Wesley, 1969.

Kiparsky, P. & Kiparsky, C. Fact. I: Steinberg, D. & Jakobovits, L.A. (Eds.) Semantics. An interdisciplinary reader in philosophy, language and psychology. Cambridge: University Press, 1971. Ss 345-369.

Klann, G. Aspekte und Probleme der linguistischen Analyse schichtenspezifischen Sprachgebrauchs. Studien und Berichte (Berlin: Max Planck-Institut für Bildungsforschung), Nr 31, 1975.

Lakoff, G. Instrumental adverbs and the concept of deep structure. Foundations of Language, 1968, 4, 4-29.

Levelt, W.J.M. Formal grammars in linguistics and psycholinguistics. III. Psycholinguistic applications. The Hague: Mouton, 1974.

Lewis, D. General semantics. I: Davidson, D. & Harman, G. Semantics of natural language. Dordrecht, Holland: Reidel, 1972. Ss 169-218.

Loman, B. Om relationen mellan ordföljd och betydelse hos framförställda attributiva bestämningar till substantiviska huvudord. Arkiv för nordisk filologi, 1956, 71, 218-244.

Loman, B. & Jörgensen, N. Manual för analys och beskrivning av makrosyntagmer. Lund: Studentlitteratur, 1971.

Lyons, J. Introduction to theoretical linguistics. Cambridge: University Press, 1971.

Oomen, U. Automatische syntaktische Analyse. The Hague: Mouton, 1968.

Osgood, C.E., Saporta, S. & Nunnally, J.C. Evaluative assertion analysis. Litera, 1956, 3, 47-102.

Perfetti, C.A. Psychosemantics: Some cognitive aspects of structural meaning. Psychological Bulletin, 1972, 78(4), 241-259.

Reid, S. Toward a grammar of the image. Psychological Bulletin, 1974, 81(6), 319-334.

Ross, J.R. Adjectives as noun phrases. I: Reibel, D.A. & Schane, S.A. (Eds.) Modern studies in English. Englewood Cliffs, N.J.: Prentice Hall, 1969.

Schank, R.C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 1972, 3(4), 552-631.

Schank, R.C. Identification of conceptualizations underlying natural language. I: Schank, R.C. & Colby, K.M. Computer models of thought and language. San Francisco: Freeman, 1973. Ss 187-247.

Schank, R.C. (Ed.) Conceptual information processing. New York: Oxford American Elsevier, 1976.

Shafto, M. The space for case. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 551-562.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: Mc Graw-Hill, 1956.

Snider, J.G. & Osgood, C.E. (Eds.) Semantic differential technique. Chicago, Ill.: Aldine, 1969.

Stone, P., Dunphy, D., Smith, M. & Ogilvie, D. The general inquirer. Cambridge: M.I.T. Press, 1966.

Stone, P. Standards for computer-aided content analysis: The Pisa conventions and recommendations. Social Science Information, 1975, 14(1), 127-137.

Tables of the cumulative binomial distribution. Cambridge, Mass.: Harvard University Press, 1955.

Teleman, U. Definita och indefinita attribut i nusvenskan. Lund: Studentlitteratur, 1969.

Teleman, U. Tre uppsatser om grammatik. Lund: Studentlitteratur, 1972.

Teleman, U. Manual för grammatisk beskrivning av talad och skriven svenska. Lund: Studentlitteratur, 1974.

af Trampe, P. & Viberg, Å. Allmän språkteori och grammatik. Lund: Gleerup, 1972.

Welin, C.W. Om konjunktiv i svenskan. PILUS (Stockholm: Institutionen för lingvistik), Nr 1, 1970.

Welin, C.W. Semantic networks and case grammar. PILUS (Stockholm: Institutionen för lingvistik), Nr 29, 1975.

11. BILAGOR

Hjälp för identifikation av koder

Bil 1:1

Satsmodifierare - adverb och adverbfraser

Ett-ordsuttryck och fraser listade enligt frekvens i SÖK-projektets intervjumaterial. Ord med frekvens under 5 fall finns under strecket i alfabetisk ordning. Ord som inte har påträffats men som kan höra till samma kategori listas därefter.

ju
väl
också
alltså
kanske
bara
nog
egentligen
naturligtvis
liksom
helt enkelt
faktiskt
över huvudtaget
ändå
verkligen
i alla fall
i och för sig
möjligen
åtminstone
i så fall (så)
nästan
även
säkert
i stort sett
annars
framför allt
förstås
i första hand
närmast
snarast
just
nämligen
på ett sätt
till och med
eventuellt

---

absolut
allmänt
antagligen
bestämt
definitivt
dessutom
däremot
exempelvis
fan, fan i mig
fortfarande
förhoppningsvis
förmodligen
förresten
först och främst
förvisso
givetvis
gärna
herregud
huvudsakligen
i alla händelser
i allmänhet
i det närmaste
i grund och botten
i huvudsak
i högsta mening
i många fall
i princip
i regel
i stort
i stället
i synnerhet
i varje fall
i vilket fall som helst
ingenting annat än
inte bara
inte mer än
inte minst
konstigt nog
likaväl
lämpligen
mycket väl
många gånger
möjligtvis
praktiskt taget
precis
på allvar
på det hela taget
på en gång
på något sätt (vis)
redan
rent av
sannerligen
självfallet
självklart
så att säga
såklart
således
säkerligen
troligen
troligtvis
trots allt
tvunget
tvärt om
tydligen
tyvärr
ungefär
vanligen
vidare
åter (också)
ännu
äntligen
ärligt talat

---

avgjort
beklagligtvis
dessvärre
för all del
gudskelov
gunås
i onödan
i själva verket
kanhända
lyckligtvis
minsann
måhända
märkligt nog
nödvändigtvis
ofrånkomligen
onödigtvis
otvivelaktigt
oåterkalleligen
relativt sett
rimligen
rimligtvis
sannolikt
sic
strängt taget
turligt nog
tvivelsutan
tydligtvis
törhända
underligt nog
uppenbarligen
utan tvivel
visst

Satsmodifierare - satser

Nedan listas direkt från datorutskrift och i alfabetisk ordning sådana satsvärdiga uttryck som påträffats i SÖK-projektets material och som kodats med 42. De ojämförligt vanligaste har strukits under.

Denna digra autentiska samling bör ge kodaren en bild av typen av uttryck.

anser jag
det får vi nog säga
det har ju visat sig att
det hör till bilden här att
det kan det innebära att
det kan jag inte påstå
det kan jag tala om att
det kan jag väl säga att
det kan ju hända att
det kan man ju inte komma ifrån
det kan man säga
det kan vi nog säga att
det kan vi väl säga
det låter så löjligt men
det låter väldigt konstigt
det skulle väl kunna tänkas att
det var väl närmast det att
det vete fan om
det vete katten
det vill jag inte säga
det är bara det att
det är det alltså att
det är det väl så att
det är fullt riktigt
det är helt klart
det är i korthet helt enkelt så att
det är ingen tvekan om att
det är jag inte klar över
det är ju alldeles givet att
det är ju alldeles uppenbart att
det är ju klart att
det är ju så att
det är ju tyvärr så att
det är klart
det är klart att
det är liksom bara att
det är liksom så att
det är mycket möjligt att
det är möjligt att
det är rättare att säga att
det är sant
det är självklart
det är självklart att
det är säkert
det är tyvärr så att
det är uppenbart så att
det är väl mest att

du (man) kan också säga att
du (man) tycker
då menar jag det att
då menar jag som så att
då är det klart att
där kan jag säga att
där är det tyvärr så att
eller vad man ska säga
enligt min mening
enligt mitt sätt att se
från min sida
får jag säga
får jag väl lov att erkänna
får vi väl säga
har jag en känsla av
ibland är det ju fruktansvärt taskigt att säga det (att)
ja det var väl
ja det är det väl
ja jag tror att
jag anser
jag anser att
jag bedömde det (som)
jag bedömer det (som)
jag får nog säga att
jag får säga
jag får väl säga att
jag har en känsla av att
jag har väl den känslan av att
jag kan föreställa mig att
jag kan inom parentes nämna
jag kan inte riktigt säga
jag kan ju säga som så att
jag kan nog säga
jag kan också tänka mig att
jag kan säga
jag kan säga att
jag kan säga så här att
jag kan väl säga så att
jag menar
jag menar alltså att
jag menar att
jag måste säga att
jag skulle möjligen säga det att
jag skulle nog vilja säga att
jag skulle tro att
jag skulle vilja säga att

jag skulle vilja säga det att
jag skulle vilja säga det också att
jag tror
jag tror alltså att
jag tror att
jag tror det
jag tror nog att
jag tror t ex att
jag tycker
jag tycker att
jag tycker det
jag tycker nog att
jag tycker själv att
jag tycker väl
jag tyckte
jag tyckte att
jag tänker just på
jag upplever
jag upplever att
jag upplever det (som)
jag vill inte säga att
jag vill säga att
just när det gäller mitt fall
kan jag inte säga att
kan jag nog säga
kan jag säga
kan jag säga att
kan jag väl säga
kan man säga att
kan man tänka sig att
kan man väl säga
kan man väl tycka på ett sätt att
kan säga så här också
kan vi säga
kanske är personligt för mig
låt oss säga
man kan säga
man kan säga att
man kan säga så att
man skulle kunna säga det att
man (jag) trodde att
med det menar jag att
med det får jag säga det att
men det är en osäker uppgift
men då var det så att
men så är det
menar jag
menar jag att
måste jag framhålla
måste jag ju säga
måste jag säga att
nu ska man säga att
nu tror jag
nu vill jag säga att
nu är det inte så att
nu är det så att
nu är det väl så att

när det kommer till kritan
och sen får jag säga det att
om du förstår
om jag får uttrycka det så
om jag talar för mig själv
om man säger så
sen får jag nog säga att
ska jag vara riktigt ärlig va
ska vi kalla det för
ska vi säga
skulle jag tro
skulle jag tycka
skulle jag vilja säga
skulle jag väl tippa att
skulle man kunna säga att
så enkelt är det
så långt det har varit möjligt
så är det ju på det sättet att
såvitt jag begriper
såvitt jag förstår
såvitt jag kan bedöma
säg att
tror jag
tror jag att
tror jag det är
tror jag också
tycker jag
tycker jag att
tycker jag inte
upplevde jag
upplever jag
upplever jag som
uppriktigt talat
vad jag förstår
vad jag menar är
vad jag menar är att
vad jag upplever
vad jag vet
vad ska jag säga
vad ska vi kalla det för
vad ska vi säga
vet du
vet jag
vi anser
vi kan säga att
vi kan väl säga
vi kan väl säga det att
vi säger väl att
vi tyckte
vi uppfattar (som)
vill jag minnas
vill jag säga
vågar jag nog påstå

Hjälp för identifikation av koder

Bil 1:3

Verb med två objekt (50 och 70)

Denna lista i alfabetisk ordning presenterar verb som kodats med två objekt i vårt autentiska intervjumaterial. Prepositioner och andra partiklar har medtagits endast vid tvetydiga verb och idiom. Listan ger ingen frekvens.

använda
applicera

basera
be
begränsa
begära
beställa
betala

diskutera

erhålla

flytta över
fråga
fylla
få
få en bild av
få in
få tips om
få ut
förelägga
förstärka
förvandla

ge
ge fria händer
gå ut med (=ung. distribuera)
göra

ha nytta av
hjälpa
hålla med
hänvisa

informera
införa

klämma
knycka
komplettera

lova
låsa (i)
lägga in
lägga över
lämna
lära

pracka
presentera

rekommendera

skaffa
skicka
skriva
stapla
stimulera
styra
ställa inför
säga

ta
tala
testa
tolka
täcka upp

underlätta
utbilda

vinna

öva
överföra (äv. föra över)

Kodningsschema

Kol	Variabel	Pos	Anm
	Identifikation		
1-2	Intervjuperson nr	01-40	
3-4	Fråga nr	01-n	
5-8	Mening nr	0000-n	
9-10	Ordenhets nr inom mening	01-n	
	Satstemakoder		se tillägg
11	Källa till utsagan	1, 2	
12	Negation	1, 2, 3	
13	Tempus	1, 2, 3	
14	Modus	1, 2, 3, 4, 5	
15	Villkor	1, 2	
16	Orsak	1, 2	
17	Medgivande	1, 2	
18	Följd eller avsikt	1, 2	
19	Disjunktion	1, 2	
20	Komparation	1	
21	Fråga	1	
22	Antagande	1	
23	Önskan	1	
24	Antal kort (rader) av ordenhet, där enheten inte fått plats på en rad	2-9	
25	Inbördes ordning hos kort (rad) som innehåller samma ordenhet)	1-9	
26-65	**Text**		
	Huvud-bisatskoder		
66-67	Huvudsatskolumner		
68	Hänvisningskolumn till bisatskolumner	2, 3, 4, 5	
69-70	Första-bisatskolumner		
71	Hänvisningskolumn till andra satskolumner	1, 3, 4, 5	
72-73	Andra bisatskolumner		
74	Hänvisningskolumn till andra satskolumner	1, 2, 4, 5	
75-76	Tredje bisatskolumner		
77	Hänvisningskolumn till andra satskolumner	1, 2, 3, 5	
78-79	Fjärde bisatskolumner		
80	Hänvisningskolumn till andra satskolumner	1, 2, 3, 4	

Tillägg

Satstemakoder

Kol		
11	Källa till utsagan	1 = anger den talande själv 2 = anger att annan än talaren står för utsagan
12	Negation	1 = inte, icke, ej, aldrig, ingen, ingenting 2 = knappast o.likn. 3 = varken (vare sig) - eller
13	Tempus	1 = nuvarande tid, gäller när utsagan görs 2 = förfluten tid från tillfället då utsagan görs 3 = kommande tid från tillfället då utsagan görs
14	Modus	1 = indikativ 2 = imperativ 3 = konjunktiv form 4 = konjunktiv betydelse 5 = vissa hjälpverb
15	Villkor	1 = villkorssatsen 2 = följdsatsen
16	Orsak	1 = orsakssatsen 2 = följdsatsen
17	Medgivande	1 = medgivandet 2 = inskränkningen
18	Följd/avsikt	1 = försatsen 2 = resultatsatsen
19	Disjunktion	1 = antingen, visserligen å ena sidan 2 = --- eller, --- men, å andra sidan
20	Komparation	1 = förekomst
21	Fråga	1 = förekomst
22	Antagande	1 = förekomst
23	Önskan	1 = förekomst

Scheman för kodning och stansning Bil 2:2

Stansinstruktion

Kol 1 - 25	Sifferkoder
Kol 1 - 2	ip nr 01-40
" 3 - 4	fråga nr 05-08
" 5 - 8	mening nr (0001-1000)
" 9 - 10	ord nr i mening 01-90
" 11 - 14	koder ges vid första ordets början
	skall stansas hela vägen ner till ny mening (kol 5-8)

Kol 12 har -, betyder blank (= ingen stansning), liksom övriga Kol 15 - 25 är blanka (med eller utan -). (Strecket (-) i kol 12 finns för kodarnas minnes skull, eftersom i denna kol kan finnas en annan siffra som är betydelsefull.)

Alltså kol 11 - 23:

Där kod ges gäller koden hela vägen ner tills den bryts av ny kod eller avbryts till blank genom -. Koderna bryts alltid vid ny mening (kol 5-8).

Kol 24 - 25

Där kod ges gäller den endast för det kort den står på. Om i kol 24 står 3 (3 ggr i följd) har tre kort samma kod i kol 9-10. I kol 25 står angivet ordningsföljden i korten. I detta fall 1, 2, 3.

Kol 26 - 65 Text

Kol 66 - 80 Sifferkoder

Dessa kol är uppdelade i "spalter".

Kol 66 - 67	= spalt 1
" 69 - 70	= " 2
" 72 - 73	= " 3
" 75 - 76	= " 4
" 78 - 79	= " 5

Kol 68, 71, 74, 77, 80 anger med siffra (1-5) att fortsättning följer i spalt nr. Kod kan också ges på samma kort.

```
Ex:  kol   666768697071        kol   66676869707172 73
     kod    3 0 2                     3 3 2 3 0
     "            5 0                       4 0
     "            4 0                       5 0
     "            3 0 1                     5 3 3 5 0
     "      4 0
     "      5 0
```

Koder som kan förekomma i kol 66-67, 69-70, 72-73, 75-76, 78-79 är:
30, 31, 32, 33, /40, 41, 42, 43, 44, 45, 46, /50, 51, 52, 53, /60, 61, 62, 63, /70, 71, 72, 73, /80, 81, 82, 83, 88, /99, /++ (plus), -- (minus).

Förkunskapstest Bil 3:1

Text

(Texten är hämtad ur Bierschenk, I. Datorbaserad innehållsanalys: Teoretiska och praktiska överväganden. Pedagogisk-psykologiska problem, Nr 283, 1975.)

I den samhällsvetenskapliga forskningen används i stor utsträckning innehållsanalys när man vill ha information ur verbala data, t ex intervjusvar, politiska tal eller tidningstexter. Det finns emellertid ingen innehållsanalysmetod som gäller generellt för allt verbalt material. Eftersom forskare inom olika discipliner sedan länge har varit bundna till manuella analyser, har det utvecklats lika många innehållsanalystekniker som det funnits forskare som använt innehållsanalysmetoder. Först nu tycks utvecklingen ha kommit så långt att denna metod kan bli mera enhetlig och därför mera generell i och med att datorer kan användas.

Förkunskapstest Bil 3:2

Uppgiftsblad

Uppgift 1.

Läst först igenom texten.

På ett löst blad som Du får finner Du fem spalter med rubriker för fem ordklasser. Skriv upp orden i den ordning de kommer i texten under den rubrik som Du tycker är den riktliga. Om något ord inte passar in, hoppar Du över det. Hoppa inte över något ord, även om det redan har förekommit.

Tid: 10 minuter

Uppgift 2.

Som Du vet byggs en text upp genom meningar och satser. Det finns meningar som har flera satser. Din uppgift här är att markera början och slut på satser på följande sätt.

a. Där en mening börjar och slutar sätter Du dubbla snedstreck //

b. Om en mening har flera satser, sätter Du ett enkelt snedstreck / mellan två satser, dvs där en sats slutar och en annan börjar.

Tid: 5 minuter

Uppgift 3.

Du har under Din skol- eller studietid lärt Dig att det är skillnad på ordklasser och satsdelar. I den här uppgiften ska Du ta ut satsdelar i samma text som i uppgifterna 1 och 2. Satsdelarna Du ska ange är: subjekt, predikat, objekt, predikatsfyllnad, attribut och adverbial.

Du ska skriva beteckningar för satsdelarna under respektive ord på följande sätt:

subjekt = s, predikat = p, predikatsfyllnad = pf, objekt = o
attribut = att, adverbial = adv

Om Du är van vid andra symboler för satsdelar, får Du använda dessa, men då måste Du ange efter texten vilka Du använt istället.

Om Du tycker att några ord hör ihop till samma satsdel sätter Du en klammer som binder ihop delarna + beteckningen mitt under. Obs! att det är nödvändigt i vissa fall.

Tid: 15 minuter

När Du löst dessa tre uppgifter lämnar Du tillbaka samtliga papper till instruktören och skriver namn eller kod på provbladen!

Förkunskapstest

Bil 3:3

Rättningsmall: Ordklasser

Rätt svar på uppgift 1: Ordklasser ges i följande uppställning:

Substantiv	verb	adjektiv	pronomen	adverb
forskningen	används	samhällsvetenskapl.	den	när
utsträckning	vill	stor	man	t ex
innehållsanalys	ha	verbala	det	emellertid
information	finns	politiska	ingen	sedan
data	gäller	generell	som	länge
intervjusvar	har	verbalt	allt	många
tal	varit	olika	det	som
tidningstexter	bundna	manuella	det	först
innehållsanalys-	har	lika	som	nu
metod	utvecklats	enhetlig	denna	så
material	funnits	generell		mera
forskare	använt			därför
discipliner	tycks			mera
analyser	ha			i och med
innehållsanalys-	kommit			
tekniker	kan			
forskare	bli			
innehållsanalys-	kan			
metoder	användas			
utvecklingen				
metod				
datorer				

Samtliga ord skall vara antecknade. Böjningsformer spelar ingen roll. Adverben 'sedan länge', 'först nu' får vara hopskrivna, men räknas som två ord. Detta gäller också verb, som sammansatts med hjälp- och huvudverb. 'som' skall vara tre sammanlagt, men ett 'som' som adverb får föras till pronomen. 'lika' och 'olika' får föras till adverb. Dessa har förts till adjektiv här, eftersom de oftast i vårt material har betraktats som adjektiv. 'ingen' får föras till adjektiv, eftersom det kan vara både och i form av "prominellt adjektiv" (se SAOL).

Rätt svar på frågan är: Totalt 73 ord, med ovan beskrivna variationer. Poängtal max 73. Fel infört ord ger 1 minus. Överhoppat ger 1 minus.

Följande ord kan tänkas vara uteslutna eller felskrivna: 'bundna' (om det har betraktats som adjektiv), 't ex'(om det har utlästs som preposition + substantiv, där då substantiven ökar med ett ord), 'i och med' (som aldrig sammanlästs och därmed inte kommer med). Summa tre fel godkänns alltså, om de är de ovan nämnda. Inte annars.

Förkunskapstest Bil 3:4

Rättningsmall: Satser

Lösning:

// I den samhällsvetenskapliga forskningen används i stor utsträckning innehållsanalys/ när man vill ha information ur verbala data, t ex intervjusvar, politiska tal eller tidningstexter. // Det finns emellertid ingen innehålls-analysmetod / som gäller generellt för allt verbalt material. // Eftersom forskare inom olika discipliner sedan länge har varit bundna till manuella analyser, / har det utvecklats lika många innehållstekniker / som det funnits forskare / som använt innehållsanalysmetoder. // Först nu tycks utvecklingen ha kommit så långt / att denna metod kan bli mera enhetlig och därför mera generell / i och med att datorer kan användas. //

Kommentar: Vi har utgått ifrån regeln "ett subjekt och ett predikat i varje sats," eftersom det är så man traditionellt brukar definiera en fullständig sats.

Fyra meningar och 11 satser.

Det kan tillåtas att de första och de sista snedstrecken utelämnats. Det kan ju inte missförstås.

Ett fel tillåts, nämligen glömt snedstreck i mening 3 vid 'innehållstekniker/ som'. Poängtal max 1 per markering. Mening 1 = 1, mening 2 = 1, mening 3 = 3, mening 4 = 2, dvs totalt 7.

Godkänt 6 p, med ovan beskrivna fel.

Skulle någon mot förmodan ha markerat med enkelt streck vid meningsslut eller med dubbelt inom mening är provet inte godkänt, oavsett antal.

Förkunskapstest Bil 3:5

Rättningsmall: Satsdelar

Lösning:

I den samhällsvetenskapliga forskningen används i stor utsträckning
adv p adv
innehållsanalys när man vill ha information ur verbala data, t ex
s adv s p o adv
intervjusvar, politiska tal eller tidningstexter. Det finns emellertid
att att att s p adv
ingen innehållsanalysmetod som gäller generellt för allt verbalt material.
att s s p adv adv
Eftersom forskare inom olika discipliner sedan länge har varit bundna
s att adv p
till manuella analyser, har det utvecklats lika många innehållsanalystekniker
adv p s p adv s
som det funnits forskare som använt innehållsanalysmetoder. Först nu
adv s p s s p o adv
tycks utvecklingen ha kommit så långt att denna metod kan bli mera enhetlig,
p s p adv s p pf
och därför mera generell i och med att datorer kan användas.
adv pf adv s p

51 är maxpoäng.

Av dessa måste samtliga "predikat" vara rätt angivna. Om 'p' angivits under varje verbform räknas det naturligtvis som rätt men 1 poäng ges för hela uttrycket. Har däremot endast huvudverbet markerats dras 1/2 poäng av.

Samtliga subjekt och objekt måste vara korrekt angivna. Tre avvikelser kan tillåtas, nämligen 3 fall av formellt subjekt ('det'), om det lämnats omarkerat. Har de däremot angivits med annan markering, räknas 1 minuspoäng per typ.

Prediktasfyllnad skall vara korrekt, men klammer behöver inte ha markerats. Däremot skall rätt ord ha angivits (huvudordet i sammansättningen).

Adverbial och attribut är ibland svåra att skilja åt. Det första 'adv' skall vara korrekt samt uttrycken 'i stor utsträckning', 'sedan länge', 'först nu', 'så långt', 'i och med att', som dessutom skall ha klamrar.
Korrekta skall också vara adv 'här', 'emellertid', 'generellt' och 'därför'.
Attributen 'ingen' och 'inom olika discipliner' skall vara korrekta.

Godkänt resultat är
s = 11, p + pf = 15, o = 2, adv = 10, att = 2. Summa 40.

Test på autentiskt material   Bil 4:1

Text

(Texten är hämtad ur intervjuundersökningens provintervjuer)

Intervjuaren: Jo, när det gäller vilka typer av problem man borde forska kring, så skulle jag gärna vilja veta om du har några typexempel som du helst skulle vilja se att man tar upp i skolforskningen.

Provtext (intervjupersonens svar):

Jag ser det nog väldigt konventionellt som gammal lärare, och jag skulle egentligen kunna formulera det med förkortningen SIA, sådana problem som har med skolans inre arbete att göra, men kanske i rätt stor utsträckning lagd på undervisningssidan. Jag vet att detta i dag inte uppfattas som särskilt radikalt eller intressant eller så där, men det har alltså under alla dom år jag har varit lärare varit ett personligt intresse att göra undervisningen så effektiv som möjligt. Jag tror att det ligger rätt mycket emotionell tillfredsställelse för både elever och lärare om dom uppfattar det vi gör som vettigt och få upplagt att dom lär sig någonting, så klarar vi nog rätt mycket av dom emotionella problemen samtidigt utan att vi behöver säga så mycket om terapi och elevvård, därför att mycket av bra elever ligger i att ge dom bra undervisning, så att det tror jag så länge jag arbetar kommer att vara ett centralt problem för min del, att förbättra undervisningssituationen.

Test på autentiskt material

Bil 4:2

Uppgiftsblad

Uppgift 1.

Läs först igenom texten.

På ett löst blad som Du får finner Du fem spalter med rubriker för fem ordklasser. Skriv upp orden i den ordning de kommer i texten under den rubrik som Du tycker är den riktiga. Om något ord inte passar in, hoppar Du över det. Obs att vissa ord tillsammans kan bilda en ordklass. Uppgiften avser halva texten, t o m 'möjligt'.

Tid: 15 minuter

Uppgift 2.

I det första testet Du fick skulle Du definiera meningar och satser. Sådant vållar inga större problem med skriven text. Den text Du har här ser annorlunda ut, eftersom den är en nedskriven talad text, som Du nog på flera ställen kan se. Den som skrivit ner texten har satt komma och punkt på några ställen, där det just för den personen verkade naturligt. Nu gäller det för dig att bortse från detta och markera var Du anser att ny mening börjar och vilka satser som finns inom varje mening. Markera på följande sätt:

a. Där en mening börjar och slutar sätter Du dubbla snedstreck

b. Om en mening har flera satser, sätter Du ett enkelt snedstreck mellan två satser, dvs där en sats slutar och en annan börjar.

Tid: 10 minuter

Uppgift 3.

I texten har några ord och ordgrupper strukits under. Ange med förkortningar vilken satsdel Du tycker att de representerar.

Tid: 5 minuter

Lämna tillbaka samtliga papper!

Test på autentiskt material | Bil 4:3

Rättningsmall: Ordklasser

Rätt svar på uppgift 1: Ordklasser ges i följande uppställning

Substantiv	verb	adjektiv	pronomen	adverb
lärare	ser	konventionell	jag	nog
förkortningen	skulle	inre	det	väldigt
SIA	kunna	gammal	som	egentligen
problem	formulera	stor	jag	kanske
skolans	har	radikalt	det	rätt
arbete	att göra	intressant	sådana	idag
utsträckning	lagd	personligt	som	inte
undervisnings-	vet	effektiv	jag	särskilt
sidan	uppfattas		det	så där
är	har		alla	alltså
lärare	har		dom	under
intresse	varit		jag	så...som
undervisningen	varit			möjligt
	göra			

Samtliga ord skall vara antecknade. Böjningsformer spelar ingen roll.

Verben får vara hopskrivna, men räknas som två ord.

Adverbet 'så där' får skrivas som två ord, men räknas som ett.

'Rätt' och 'möjligt' får räknas till adjektiv.

'Inre' kan få räknas till adverb eller pronomen.

Rätt svar på frågan är: Totalt 60 ord, med ovan beskrivna variationer.

Poängtal max 60. Fel infört ord ger 1 minus. Överhoppat ger 1 minus.

Test på autentiskt material Bil 4:4

Rättningsmall: Satser

Lösning:

// Jag ser det nog väldigt konventionellt som gammal lärare // Och jag skulle egentligen kunna formulera det med förkortningen SIA, / sådana problem som har med skolans inre arbete att göra / men kanske i rätt stor utsträckning lagd på undervisningssidan. // Jag vet att detta idag inte uppfattas som särskilt radikalt eller intressant eller så där, // men det har alltså under alla dom år / jag har varit lärare / varit ett personligt intresse / att göra undervisningen så effektiv som möjligt. // Jag tror / att det ligger rätt mycket emotionell tillfredsställelse för både elever och lärare / om dom uppfattar det / vi gör / som vettigt // och få upplagt / att dom lär sig någonting, // så klarar vi nog rätt mycket av dom emotionella problemen samtidigt/ utan att vi behöver säga så mycket om terapi och elevvård, // därför att mycket av bra elever ligger i att ge dom bra undervisning, // så att det / tror jag / så länge jag arbetar / kommer att vara ett centralt problem för min del / att förbättra undervisningssiutationen. //

Antal markeringar och maxpoäng är 25.

Här kan flera lösningar tänkas. Det är det som är problematiskt med taltexter.

Följande markeringar skall vara korrekta

// Jag, lärare // Och, SIA, / sådana

undervisningssidan. // Jag, så där, // men, år / jag

intresse / att, möjligt. // Jag, lärare / om, det / vi (med alternativet uppfattar / det)

vettigt // och, upplagt / att, någonting, // så, samtidigt / utan,

elevvård, // därför, undervisning, // så, del / att

Därav kan följande avvikelser tänkas:

lärare // om, vettigt / och, någonting, / så

(Det går att uppfatta meningen så istället.)

Godkänt är alltså 16 poäng.

Test på autentiskt material Bil 4:5

Rättningsmall: Satsdelar

Lösning:

kunna formulera ... det ... med förkortningen SIA ... i dag ...
p o adv adv

som särskilt radikalt ... ett personligt intresse ... att göra under-
pf pf s

visningen så effektiv som möjligt ... tillfredsställelse ... dom ...
s s s

någonting ... bra undervisning ... så länge jag arbetar ...
o o adv

ett centralt problem .... .
pf

13 är maxpoäng.

Fel tillåts vid det första pf 'som särskilt radikalt', men inte vid det andra och tredje, där nyckelverbets 'vara' bör vägleda.

Subjektsatsen (att...) är svår, varför fel tillåts, liksom subjektet 'tillfredsställelse'.

Tre avvikelser får alltså förekomma. Godkänt är 10 poäng.

Anvisningar för markering av överensstämmelse mellan två kodare

Ruta 1. Beteckningar vid jämförelse av begreppsavgränsning

Beteck-ning	Överensstämmelse/avvikelsetyp	Markeras för Kodare A	Kodare B
	Överensstämmelse mellan kodare.		
	Samma begrepp innehåller samma ord hos båda kodarna.		
AO, +A	Avvikelser vid begreppsavgränsning		
	Kodare A har avgränsat ett begrepp annorlunda än kodare B, så att ett ord mer eller mindre ingår i samma begrepp.	AO	AO
	Kodare A har ett extra begrepp pga annorlunda avgränsning än kodare B.	+A (kodare A)	ingen markering
	Kodare B har ett extra begrepp pga annorlunda avgränsning än kodare A.	ingen markering	+A (kodare B)
	Överensstämmelse vid supplering.*		
	Kodare A och B har supplerat samma ord på samma sätt, så att samma begrepp innehåller samma ord hos båda kodarna.		
-S, SO, +S	Avvikelser vid supplering		
	Kodare A har supplerat ett ord som kodare B ej har supplerat.	-S	-S
	Kodare A och B har båda supplerat ett ord, men på olika sätt.	SO	SO
	Kodare A har insupplerat ett begrepp som kodare B inte har.	+S (kodare A)	ingen markering
	Kodare B har insupplerat ett begrepp som kodare A inte har.	ingen markering	+S (kodare B)
	Överensstämmelse vid strykning		
	Samma begrepp har strukits av båda kodarna. Begreppet utgår ur analysen och uppförs ej på datablad.	ingen markering	ingen markering
+	Avvikelse vid strykning		
	Kodare A har strukit ett begrepp som kodare B inte har strukit; kodare B har extra begrepp	ingen markering	+ (kodare B)
	Kodare B har strukit ett begrepp som kodare A inte har strukit; kodare A har extra begrepp.	+ (kodare A)	ingen markering

* Beteckningen för överensstämmelse vid supplering är densamma som vid överensstämmelse mellan kodare. Det rör sig dock om två olika typer av överensstämmelse. Dessa två typer har emellertid behandlats tillsammans och de kan ej urskiljas i materialet.

Ruta 2. Jämförelse av satsavgränsning

Satsen bedöms av som	Kodare A	Kodare B	Förklaring	Beteckning
Huvudsats	1	1	Samma sats kodas som huvudsats	Gemensam huvudsats
Bisats	1	1	Samma sats kodas som bisats	Gemensam bisats
Huvudsats Bisats	1 -	- 1	Samma sats kodas som huvudsats av kodare A och som bisats av kodare B	Avvikande huvudsats (kodare A) Avvikande bisats (kodare B)
Huvudsats	1	-	Kodare A har urskiljt en extra huvudsats	Avvikande (extra) huvudsats (kodare A)
Bisats	1	-	Kodare A har urskiljt en extra bisats	Avvikande (extra) bisats (kodare A)

Ruta 3. Jämförelse av innehållskodning av begrepp

Beteckning	Överensstämmelse/avvikelsetyp	Markeras för Kodare A	Markeras för Kodare B
	Överensstämmelse mellan kodare		
	Samma begrepp har tilldelats samma innehållskod		
-x	Gemensam avvikelse mellan kodare		
	Samma begrepp har tilldelats olika innehållskoder	-x	-x
-.	Överensstämmelse mellan kodare		
	Samma beroende begrepp har tilldelats samma attributkod men olika huvudkod. Avvikande huvudkod har redan markerats som gemensam avvikelse vid huvudordet.	-.	-.
	Enskild avvikelse		
-	Kodare A har extra begrepp i texten och har därför extra innehållskoder.	-	ingen markering
	Kodare B har extra begrepp i texten och har därför extra innehållskoder.	ingen markering	-

Ruta 4. Jämförelse av kodning av satstema

Beteck-ning	Överensstämmelse/avvikelsetyp	Markeras för Kodare A	Kodare B
	Överensstämmelse mellan kodare		
	Samma sats har tilldelats samma satskoder för källa, tempus, modus (ktm) av kodarna	.	.
	Samma sats har tilldelats samma satskoder för "övriga satstema" av kodarna	.	.
-x	Gemensam avvikelse		
	Samma sats har tilldelats olika satskoder för ktm av kodarna	x-	x-
	Endast kodare A har kodat satsens ktm eller negation i samma sats	x-	x-
	I samma sats har tilldelats olika satskoder för "övriga satstema" av kodarna. Endast en avvikelsemarkering uppförs för varje kodare, även om två koder krävs för satstemat. Kodarna var överens om att ett "övrigt" satstema förekom i satsen, men kodade temat olika.	x-	x-
-	Enskild avvikelse		
	Avvikelser vid kodning av ktm och "övriga satstema", som orsakas av olikheter vid avgränsning av begrepp eller satser markeras för varje kodare som enskild avvikelse.	-	-
	Kodare A har supplerat in en extra sats och uppfört satskoder.	-	ingen markering
	Endast kodare A har identifierat ett satstema och uppfört satskod.	-	ingen markering
	Vid kodning av "övriga satstema" där två markeringar krävs, har kodare A endast uppfört en markering.	-	ingen markering
	Vid kodning av "övriga satstema" där två markeringar krävs, har kodare B endast uppfört en markering.	ingen markering	-

Anm: Rutorna 1-4 i denna bilaga är hämtade ur Berg, M. Reliabilitetsprövning av en metod för innehållsanalys av intervjutext. Testkonstruktion och testdata. Nr 26, 1974.

Den innehåller dessutom en diskussion om överensstämmelser och redovisar resultat från användningen av den första versionen av manualen:

Bierschenk, I. Konstruktion av ett regelsystem för en datorbaserad innehållsanalys av intervjutext: Preliminärmanual och några utprövningsresultat. Testkonstruktion och testdata, Nr 25, 1974.

Detta termregister är avsett att vara till hjälp vid uppsökning av kodningsreglerna i kapitlen 4-6. Till exempelsamlingen (kap 7) görs hänvisning endast till kommenterade ställen. Siffrorna avser sida.

ISSN 0346-5039